



Pedro Ricardo Morgado Henriques

Licenciado em Bioquímica

Estudo de interações proteína-carbohidrato em sistemas bacterianos

Dissertação para obtenção do Grau de Mestre em
Biotecnologia

Orientador: Eurico J. Cabrita, Professor auxiliar, FCT-UNL
Co-orientador: Isabel de Sá-Nogueira, Professora associada
com agregação, FCT-UNL

Juri:

President: Prof. Dr. Pedro Miguel Ribeiro Viana Baptista
Arguente: Prof. Dr. Maria dos Anjos Lopez de Macedo
Vogal: Prof. Dr. Eurico José da Silva Cabrita



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro de 2014



Pedro Ricardo Morgado Henriques

Licenciado em Bioquímica

Estudo de interações proteína-carbohidrato em sistemas bacterianos

Dissertação para obtenção do Grau de Mestre em
Biotecnologia

Orientador: Eurico J. Cabrita, Professor auxiliar, FCT-UNL
Co-orientador: Isabel de Sá-Nogueira, Professora associada
com agregação, FCT-UNL

Juri:

President: Prof. Dr. Pedro Miguel Ribeiro Viana Baptista
Arguente: Prof. Dr. Maria dos Anjos Lopez de Macedo
Vogal: Prof. Dr. Eurico José da Silva Cabrita



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro de 2014

Copyright

Estudo de interações proteína-carbohidrato em sistemas bacterianos.

Pedro Ricardo Morgado Henriques, FCT-UNL, UNL

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objectivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

Agradecimentos

Gostaria de agradecer aos meus orientadores Prof. Eurico Cabrita e Prof. Isabel Sá-Nogueira, por me permitirem desenvolver a minha tese de mestrado em ambos os laboratórios.

Os meus agradecimentos estendem-se também aos meus colegas de laboratório Dr. Ângelo Figueiredo, Dr. Filipa Marcelo, Dr. Marta Corvo, Dr. Ana Sofia, Mário Ferreira e Lia Godinho, por todas as discussões construtivas. Não esquecendo os juniores quero também agradecer ao Jonhy Cascão, Micael Silva, João Silva, Ricardo Martinho e Tiago Paiva.

Em último lugar, e mais importante que tudo, quero expressar a minha gratidão à minha família e amigos.

Este trabalho foi parcialmente financiado pela Fundação para a Ciência e a Tecnologia através dos projetos: PEst-OE/BIA/UI0457/2011 e através da rede nacional de RMN (RNRMN) pelo projeto RECI/BBB-BQB/0230/2012.

Resumo

Bacillus subtilis tem a capacidade de utilizar arabinoligossacáridos presentes na parede das células vegetais, através de um consórcio de enzimas envolvidas na hidrólise enzimática dos mesmos. A captação destes carboidratos por parte do *Bacillus subtilis* depende de dois transportadores membranares AraE e AraNPQ-MsmX. Após a captação de *L*-arabinose e arabinoligossacáridos, estes substratos serão metabolizados pelos produtos enzimáticos codificados por genes pertencentes ao operão responsável pela metabolização de arabinose *araABDLMNPQ-abfA*.

Neste operão, para além dos genes responsáveis pelo transportador AraNPQ, o gene *araL* determina a síntese de uma enzima com atividade de fosfatase, AraL, que tem putativamente como principal função a destoxificação de intermediários metabólicos fosforilados, em situações particulares.

O objectivo deste trabalho consistiu em descobrir quais os determinantes moleculares envolvidos em ambos os processos de reconhecimento molecular, ora o reconhecimento de arabinoligossacáridos por parte da AraN, do ponto de vista dos carboidratos ora de carboidratos fosforilados do ponto de vista da AraL. Para elucidar estes processos de reconhecimento molecular proteína-carboidrato foi utilizada a técnica de RMN, modelação computacional e mutagenese dirigida.

No primeiro capítulo são introduzidos conceitos fundamentais para a percepção da ação de ambas as proteínas (AraN e AraL), no organismo *Bacillus subtilis*.

O segundo capítulo, refere-se ao estudo dos mecanismos envolvidos no reconhecimento proteína-carboidrato através da técnica de STD-RMN para estudar interações do ponto de vista do carboidrato, bem como abordagens bioinformáticas tais como alinhamentos de sequência primária e *dockings* moleculares, para identificar resíduos do ponto de vista da proteína passíveis de se encontrarem envolvidos no reconhecimento molecular, que numa última instância são mutados para se confirmar a sua relevância no processo de reconhecimento molecular do ponto de vista da proteína. Neste capítulo é demonstrado que a AraN é responsável pelo reconhecimento de arabinoligossacáridos e celoligossacáridos. No que diz respeito aos primeiros, a AraN reconhece preferencialmente arabinoligossacáridos com três subunidades, verificando-se a perda de saturação na subunidade não redutora da arabinotetraose. O processo de reconhecimento molecular apresenta uma constante de dissociação com uma ordem de grandeza na ordem dos μM . Do ponto de vista da proteína são identificados resíduos (W253, Y254 e Y54) putativos de se encontrarem também eles envolvidos no reconhecimento molecular do ponto de vista da AraN através de interações *CH-stacking*.

Finalmente no terceiro capítulo, é apresentada a optimização da sobreexpressão da AraL, dado que numa primeira fase a sua produção ocorria sob a forma de corpos de inclusão. Para tal são descritas metodologias de solubilização e renaturação da mesmas, com recurso a agentes caotrópicos (ureia e GdmCl) e detergentes (SDS).

Palavras-chave: STD-RMN, interação proteína-carboidrato, AraN, AraL.

Abstract

Bacillus subtilis has the ability to utilize arabinoligosaccharides present in the plant cell wall, mainly through the enzymatic hydrolysis of the former by the action of a consortium of enzymes. The acquisition of these carbohydrates from *Bacillus subtilis* standpoint depends on two transmembranar transporters AraE and AraNPQ-MsmX. After the uptake of L-arabinose and other arabinoligosaccharides, these substrates will be metabolized by the enzymatic products codified within the arabinose metabolizing operon *araABDLMNPQ-abfA*.

Within this operon, besides the genes responsible for the AraNPQ transporter, there is a gene *araL* that determines the synthesis of a phosphatase (AraL), putatively having the main function of detoxification of phosphorylated metabolic intermediates, in particular situations.

The aim of this work was to find out which molecular determinants are involved in both processes of molecular recognition, either from the ligand viewpoint in the interaction of arabinoligosaccharide-AraN, or from the protein viewpoint within the interaction of phosphorylated carbohydrates with respect to AraL. To elucidate these biomolecular processes of protein-carbohydrate recognition techniques such as NMR, computer modelling and directed mutagenesis, were used.

In the first chapter the fundamental concepts of the action of both proteins (AraN and AraL) in the organism *Bacillus subtilis* are introduced.

The second chapter, refers to the study of the mechanisms involved in protein-carbohydrate recognition by STD-NMR, with the aim of understanding the interactions from the carbohydrate viewpoint, as well as bioinformatics approaches such as sequence alignments and molecular dockings, to identify residues from the protein point of view that will be object of mutagenesis to confirm their relevance on the molecular recognition process. In this chapter it's shown that AraN has the ability of recognizing arabinoligosaccharides as well as celloligosaccharides. With regard to arabinoligosaccharides, AraN will preferably recognize arabinoligosaccharides with three monomers, being the case of saturation loss in the non-reducing end in arabinotetraose. The recognition process has a dissociation constant in the micromolar range. From the protein standpoint, residues W253, Y254 and Y54 are identified as putatively being involved in molecular recognition process through CH-*stacking* interactions.

Finally in the third chapter, there will be presented the methodologies used for the optimization of AraL overexpression, since this protein was overexpressed in a first instance in inclusion bodies. For that matter the protein in study was solubilized and refolded resorting to chaotropic agents such as urea and guanidine hydrochloride or detergent like SDS.

Key-words: STD-NMR, protein-carbohydrate interactions, AraN, AraL.

Índice

Copyright	I
Agradecimentos	III
Resumo	V
Abstract.....	VII
Índice	IX
Índice de figuras	XIII
Índice de tabelas	XV
Abreviaturas	XIX
Capítulo 1: Introdução	1
I.1. <i>L</i> -Arabinose e a parede celular	2
I.2. Transporte de arabinoligossacáridos para o interior da célula.....	4
I.2.1. Transportadores do tipo ABC.....	5
I.2.1.1. Organização estrutural dos transportadores do tipo ABC	5
I.3. Proteínas de ligação ao substrato- SBP	6
I.3.1. AraN- Proteína de reconhecimento de arabinoligossacáridos	7
I.3.2. Topologia estrutural de SBP.....	8
I.3.3. Alteração conformacional das SBP's e translocação de substrato	9
I.3.4. Especificidade das SBP's	10
I.4. Desalogenase de haloalcanos	12
I.4.1 AraL- Fosfatase de carboidratos fosforilados	13
I.5. A Ressonância Magnética Nuclear no estudo da interação proteína-ligando	13
I.5.1. STD-RMN- Transferência de saturação	19
Capítulo II: Reconhecimento molecular de arabinoligómeros pela AraN.....	21
II.1. Introdução.....	21
II.2. Materiais e métodos.....	22
II.2.1. Expressão heteróloga da AraN-His ₆ tag	23
II.2.2. Isolamento da AraN-His ₆ Tag.....	23
II.2.3. Purificação da AraN-His ₆ tag	23
II.2.4. Quantificação da AraN-His ₆ tag	24
II.2.5. Preparação dos carboidratos	25
II.2.6. Aquisição de espectros de RMN	25
II.2.6.1. Caracterização de carboidratos.....	25
II.2.6.2. Estudos de STD-NMR	26

II.2.7. NOE e trNOE	28
II.2.8. Modelo de homologia estrutural 3D da AraN	28
II.2.9. Docking molecular	28
II.2.10. Análise de homologia da sequência da AraN com outras SBP	29
II.3. Resultados e discussão	31
II.3.1. Expressão heteróloga e purificação da AraN	31
II.3.2. Caracterização de carboidratos	32
II.3.2.1. Atribuição de ressonâncias espectrais de carboidratos	32
II.3.3. Estudos de STD-RMN	39
II.3.3.1. Otimização do tempo de saturação e mapeamento de epítomos	39
II.3.3.2. Influência da temperatura	44
II.3.3.3. Cálculo de K_d	47
II.3.3.4. Estudo das interações ligando-AraN com arabinoligossacáridos	49
II.3.3.5. Estudo das interações ligando-AraN com celoligossacáridos	56
II.3.3.6. <i>Screening</i> de outros carboidratos	59
II.3.3.7. Interação em condições de competição: arabinotriose/maltohexaose e arabinotriose/celotetraose	61
II.3.4. Construção de modelos 3D para análise e interpretação de interações	65
II.3.4.1. Modelo de homologia estrutural 3D da AraN	65
II.3.5. Localização do local de ligação da AraN por alinhamento estrutural	67
II.3.6. <i>Docking</i> molecular	71
II.3.6.1. <i>Clusters</i> obtidos através do AutoDock	72
II.3.6.2. <i>Docking</i> molecular entre AraN-Arabinotriose	73
II.3.7. Mutagênese	76
II.3.8. Conformação da arabinotriose no local de ligação da AraN-estudo de NOE de transferência	78
II.4. Conclusões	80
Capítulo III: Otimização da sobreexpressão da AraL	83
III.3.1. Introdução	83
III.3.2. Materiais e métodos	83
III.3.2.1. Sobreexpressão da AraL-His ₆ Tag	85
III.3.2.2. Isolamento e purificação da AraL-His ₆ Tag	85
III.3.2.2.1. Teste de sobreexpressão à escala piloto	85
III.3.2.2.2. <i>Screening</i> de condições de sobreexpressão	85
III.3.2.2.3. Solubilização de corpos de inclusão com ureia e renaturação na coluna Ni-NTA	86
III.3.2.2.4. Solubilização de corpos de inclusão com guanidino de hidróclorato e renaturação na coluna Ni-NTA	87
III.3.2.2.5. Solubilização de corpos de inclusão com SDS	87
III.3.2.3. Verificação de atividade fosfatase da AraL	87

III.3.2.4. Monotorização do processo de solubilização e renaturação da AraL com recurso a SDS por RMN.....	88
III.3.3. Resultados e discussão	88
III.3.3.1. Solubilização e renaturação de corpos de inclusão	89
III.3.3.1.1. Teste de sobreexpressão da AraL	90
III.3.3.1.2. <i>Screening</i> de tampões de solubilização da AraL	91
III.3.3.1.3. Solubilização e renaturação da AraL com recurso a SDS	92
III.3.3.1.4. Solubilização e renaturação da AraL com recurso a ureia	98
III.3.3.1.5. Solubilização e renaturação da AraL com recurso a GmdCl.....	98
III.4. Conclusões	101
IV. Conclusões gerais e perspectivas futuras.....	101
V. Bibliografia	107
VI. Apêndice.....	117
VI.1. Fundamentos teóricos da técnica de RMN	117
VI.1.1 Desvio químico.....	117
VI.1.2. Relaxação longitudinal T_1 e relaxação transversal T_2	117
VI.1.3. Efeito Nuclear de Overhauser (NOE).....	119
VI.1.4. Movimento molecular e relaxação de spin.....	121
VI.1.5. Tempo de correlação e densidade espectral	122
VI.2. Meios de cultura	123
VI.3. Quantificação pelo método de Bradford	124
VI.4 Dados de NMR	125
VI.5 Espectros suplementares	126

Índice de figuras

Figura I.1. Interação proteína-carbohidrato.....	1
Figura I.2. Estrutura do arabinano.....	2
Figura I.3. Transporte primário e secundário de arabinoligossacáridos em <i>Bacillus subtilis</i>	3
Figura I.4. Transportador do tipo ABC.....	5
Figura I.5. Representação da SBP Malx de <i>Streptococcus pneumoniae</i>	8
Figura I.6. Mecanismo de hidrólise de fosfatos.....	13
Figura I.7. Interação do ponto de vista da proteína.....	15
Figura I.8. Espectro SEA-TROSY.....	16
Figura I.9. Movimento translacional de diferentes moléculas em solução.....	17
Figura I.10. Estudo da interação proteína-carbohidrato, por trNOE.....	18
Figura I.11. Experiência de STD-RMN.....	19
Figura II.1. Estratégia utilizada para estudar a interação protein-carbohidrato promovida por parte da AraN.....	22
Figura II.2. <i>Gridbox</i> utilizada no <i>docking</i> molecular.....	29
Figura II.3. SDS-PAGE (12,5%), com as fracções de purificação da AraN.....	31
Figura II.4. SDS-PAGE (12,5%), com as fracções de purificação da AraN.....	31
Figura II.5. Epímeros de arabinobiose.....	32
Figura II.6. Atribuição espectral da arabinobiose.....	33
Figura II.7. Estrutura e atribuição do espectro de ^1H -RMN da arabinobiose.....	34
Figura II.8. Estrutura e atribuição do espectro de ^1H -RMN da arabinotriose.....	35
Figura II.9. Estrutura e atribuição do espectro de ^1H -RMN da arabinotetraose.....	36
Figura II.10. Estrutura e atribuição do espectro de ^1H -RMN da celobiose.....	37
Figura II.11. Estrutura e atribuição do espectro de ^1H -RMN da celotetraose.....	38
Figura II.12. Factor de amplificação (A_{STD}) em função do tempo de saturação (T_{sat}).....	41
Figura II.13. STD_{fit} de cada protão.....	42
Figura II.14. Efeito da temperatura no espectro de STD-RMN.....	45
Figura II.15. Variação da densidade espectral com o movimento molecular em função da frequência de Larmor.....	46
Figura II.16. Efeito do tempo de correlação no tempo de relaxação T_1 e T_2	46
Figura II.17. Factor de amplificação (A_{STD}) em função da concentração de ligando ($[L]$).....	48
Figura II.18. Espectro de STD-RMN da arabinose com AraN.....	49
Figura II.19. Espectro de STD-RMN da arabinobiose com AraN.....	50
Figura II.20. Espectro de STD-RMN da arabinotriose com AraN.....	52
Figura II.21. Espectro de STD-RMN da arabinotetraose com AraN.....	53
Figura II.22. Espectro de STD-RMN da arabinopentose com AraN.....	55
Figura II.23. Estrutura da celobiose e celotetraose.....	56
Figura II.24. Espectro de STD-RMN da celobiose com AraN.....	57
Figura II.25. Espectro de STD-RMN da celotetraose com AraN.....	58
Figura II.26. Estrutura da maltohexaose, xilopentose e manotriose respectivamente.....	59

Figura II.27. Espectro de STD-RMN da maltohexaose com AraN.....	60
Figura II.28. Espectro de STD-RMN da xilopentaose com AraN.....	60
Figura II.29. Espectro de STD-RMN da manotriose com AraN.....	61
Figura II.30. Espectro de STD-RMN em condições de competição da arabinotriose/manotriose com AraN.....	62
Figura II.31. Espectro de STD-RMN em condições de competição da arabinotriose/celotetraose com AraN.....	63
Figura II.32. Espectro de STD-RMN em condições de competição da arabinotriose/manotriose com AraN.....	64
Figura II.33. Representação da AraN obtida no servidor I-TASSER.....	67
Figura II.34. Alinhamento das sequências primárias com recurso ao Expresso.....	69
Figura II.35. Alinhamento das sequencias primárias com recurso ao Expresso.....	70
Figura II.36. Alinhamento da estrutura da AraN com a MBP.....	70
Figura II.37. Alinhamento da estrutura da AraN com a <i>Tvu</i> CMBP-G4.....	71
Figura II.38. Conformações obtidas no <i>docking</i> molecular.....	74
Figura II.39. Mapeamento de interacções.....	74
Figura II.40. SDS-PAGE (12,5%), com as fracções de purificação da AraN(R81A).....	76
Figura II.41. SDS-PAGE (12,5%), com as fracções de purificação da AraN(R81A).....	77
Figura II.42. Espectro de STD-RMN da arabinotriose com AraN(wt) ou AraN(R81).....	78
Figura II.43. Perturbação dos desvios químicos ^1H da região anomérica.....	79
Figura II.44. Espectro 2D-NOESY, da arabinotriose livre sem AraN.....	80
Figura II.45. Espectro 2D-NOESY, da arabinotriose na presença de AraN num excesso de 1:10.....	80
Figura III.1. Estratégia utilizada para estudar a interacção protein-carbohidrato promovida por parte da AraL.....	83
Figura III.2. <i>Landscape</i> dos estados de enrolamento da proteína em função da concentração de agente desnaturante utilizado.....	90
Figura III.3. SDS-PAGE (12,5%), com as fracções de purificação da AraL.....	91
Figura III.4. SDS-PAGE (12,5%), com as fracções de purificação da AraL.....	91
Figura III.5. SDS-PAGE (12,5%), com as fracções de purificação da AraL.....	92
Figura III.6. Espectro ^1H dos diferentes sobrenadantes.....	93
Figura III.7. Espectro ^1H (com supressão de água- zgesgp) dos diferentes sobrenadantes.....	93
Figura III.8. SDS-PAGE (12,5%), com as fracções de purificação da AraL.....	94
Figura III.9. SDS-PAGE (12,5%), com as fracções de purificação da AraL.....	94
Figura III.10. Espectro ^1H da fração insolúvel.....	95
Figura III.11. Espectro ^1H da fração solúvel resultante do passo de precipitação do SDS.....	95
Figura II. 12. SDS-PAGE (12,5%), com as fracções de purificação da AraL.....	96
Figura III.13. SDS-PAGE (12,5%), com as fracções de purificação da AraL.....	96
Figura III.14. SDS-PAGE (12,5%), com as fracções de purificação da AraL.....	97
Figura III.15. SDS-PAGE (12,5%), com as fracções de purificação da AraL.....	97
Figura III.16. SDS-PAGE (12,5%), com as fracções de purificação da AraL.....	98

Figura III.17. SDS-PAGE (12,5%), com as fracções de purificação da AraL.....	99
Figura VI.1. Distribuição dos spins nucleares em equilíbrio térmico de Boltzman.....	118
Figura VI.2. Efeito NOE.....	120
Figura VI.3. Manipulação de spins nucleares após aplicação de rf em ressonância.....	121
Figura VI.4. Função densidade espectral para diferentes regimes de movimento molecular.....	122
Figura VI.5. Recta de calibração do método de Bradford.....	124
Figura VI.6. Espectro de STD-RMN em condições de competição da arabinotriose/celotetraose com AraN.....	125
Figura VI.7. Espectro de STD-RMN em condições de competição da arabinotriose/celotetraose com AraN.....	125

Índice de tabelas

Tabela I.1. Topologia de transporte.....	4
Tabela I.2. Taxas de crescimento de diferentes estirpes de <i>B. subtilis</i>	8
Tabela I.3. Classificação das SBP's.....	9
Tabela I.4. Topologia de enzimas DHA.....	12
Tabela I.5. Motivos estruturais envolvidos na hidrólise do éster de fosfato.....	12
Tabela I.6. Metodologias estabelecidas no estudo interação proteína-ligando.....	18
Tabela II.1. Titulação AraN-arabinotriose.....	27
Tabela II.2. Desvios químicos ^1H da arabinobiose.....	34
Tabela II.3. Desvios químicos ^{13}C da arabinobiose.....	34
Tabela II.4. Desvios químicos ^1H da arabinotriose.....	35
Tabela II.5. Desvios químicos ^{13}C da arabinotriose.....	35
Tabela II.6. Desvios químicos ^1H da arabinotetraose.....	36
Tabela II.7. Desvios químicos ^{13}C da arabinotetraose.....	36
Tabela II.8. Desvios químicos ^1H da celobiose.....	37
Tabela II.9. Desvios químicos ^{13}C da celobiose.....	37
Tabela II.10. Desvios químico ^1H da celotetraose.....	38
Tabela II.11. Desvios químicos ^{13}C da celotetraose.....	38
Tabela II.12. Mapeamento de epítomos para diferentes T_{sat}	43
Tabela II.13. Factor de amplificação (A_{STD}) em função da temperatura.....	47
Tabela II.14. Tabela com as pontuações obtidas do modelo de homologia da AraN.....	65
Tabela II.15. Parâmetros retirados da simulação efectuada no AutoDock.....	73
Tabela III.1. Medição de atividade fosfatase da AraL.....	98
Tabela III.2. Medição de atividade fosfatase da AraL.....	99
Tabela VI.1. Preparação de meio Luria-Bertani (LB).....	123
Tabela VI.2. Preparação de meio mínimo M9.....	123
Tabela VI.3. Preparação de solução de sais M9.....	123
Tabela VI.4. Preparação de meio auto indução.....	123
Tabela VI.5. Tabela para a preparação de amostras para a construção da recta de calibração.....	124
Tabela VI.6. Dados experimentais.....	125

Abreviaturas

Abf2	α -L-Arabinofuranosidase em <i>Bacillus subtilis</i>
AbfA	α -L-Arabinofuranosidase em <i>Bacillus subtilis</i>
Abn2	α -1,5-arabinanase em <i>Bacillus subtilis</i>
AbnA	α -1,5-arabinanase em <i>Bacillus subtilis</i>
ABP	<i>Arabinose binding protein</i>
Abs	Absorvância
ADH	Haloalcanoato desalogenase
AraE	Permease em <i>Bacillus subtilis</i>
AraL	Fosfatase do tipo ADH em <i>Bacillus subtilis</i>
AraN	Proteína de reconhecimento de substratos (SBP) em <i>Bacillus subtilis</i>
A_{STD}	Factor de amplificação de STD
B₀	Campo magnético externo
B_{ef}	Campo magnético efectivo
B_{loc}	Campo magnético local
CBM	<i>Carbohydrate binding module</i>
COSY	<i>Correlation spectroscopy</i>
D₂O	Água deuterada
DppA	Proteína de reconhecimento de dipéptidos
<i>E. coli</i>	<i>Escherichia coli</i>
GBP	<i>Galacto-N-biose binding protein</i>
GEM	<i>Goup epitope mapping</i>
GH43	Hidrólase glicosídicas da família 43
GH51	Hidrólase glicosídicas da família 51
GmdCL	Guanidino de hidrocloreto
HBP	Proteína de reconhecimento de histidinas
HIV	Vírus da imunodeficiência humana
HSQC	<i>Heteronuclear Single Quantum Coherence</i>
I₀	Intensidade do espectro de referência fora de ressonância
IPTG	Isopropilo β -D-1-tiogalactopiranosose
I_{SAT}	Intensidade do espectro em ressonância
I_{STD}	Intensidade do espectro STD
K_a	Constante de associação
K_d	Constante de dissociação
LAOBP	<i>Lysine/arginine/ornithine binding protein</i>
LB	Meio de crescimento Luria-Bertani
LBP	<i>Lacto-N-biose binding protein</i>
LIVBP	<i>Leucine/isoleucine/valine binding protein</i>
MalIEFGK₂	Transportador do tipo ABC de maltoligossacáridos
MBP	<i>Maltose Binding Protein</i>
MM	Marcador de pesos moleculares
MSA	Multiple sequence alignment
MSMx	Proteína NBD

Mw	Peso molecular
NaCl	Cloreto de Sódio
NBD	<i>Nucleotide binding domain</i>
NOE	Efeito nuclear de Overhauser
NOESY	<i>Nuclear Overhauser Effect Spectroscopy</i>
NTA	Ácido nitriloacético
O/N	<i>Overnight</i>
OppA	Proteína de reconhecimento de oligopéptidos
PDB	<i>Protein data bank</i>
PMSF	<i>Phenylmethanesulfonyl fluoride</i>
Rf	Radiofrequência
RMN	Ressonância magnética nuclear
RMSD	<i>Root-mean-square deviation</i>
SBP	<i>Substrate binding protein</i>
SDS	Sódio dodecil sulfato
SEA-TROSY	<i>Solvent exposed amides- transversal relaxation optimized spectroscopy</i>
SLC	<i>Solute carrier</i>
STD	<i>Saturation transfer difference</i>
STD_{fit}	Factor de amplificação STD em condições de velocidade inicial
T1	Relaxação longitudinal
T2	Relaxação transversal
TEMPO	2,3,4,6-tetrametil-piperidina-1-oxil
TMD	<i>Transmembranar domain</i>
TOCSY	<i>Total correlation spectroscopy</i>
trNOE	Transferência de NOE
TROSY	<i>Transversal relaxation optimized spectroscopy</i>
T_{sat}	Tempo de saturação
TvuCMBP	<i>Thermoactinomyces vulgaris</i> cyclo/maltodextrin binding protein
σ	Constante de blindagem

Capítulo 1: Introdução

Os carboidratos representam uma das quatro bio-macromoléculas presentes nos seres vivos [1]. Tal como as proteínas e o DNA estes podem formar biopolímeros através do acoplamento de monómeros das unidades constituintes, no entanto o que diferencia os carboidratos das anteriores é o facto das unidades constituintes (monossacáridos) apresentarem um conjunto de grupos funcionais passíveis de serem utilizados no processo de polimerização e ramificação, o que em última análise irá proporcionar um leque quase infinito de variações estruturais [2].

Eventos de reconhecimento molecular são essenciais numa vasta gama de processos biológicos evidenciados na natureza e representados na figura I.1. A interação proteína-carboidrato é um processo central no i) recrutamento de leucócitos nas células endoteliais em processos inflamatórios [3] , ii) mediação do processo de reconhecimento entre agentes infecciosos (bactérias e vírus patogénicos) e células alvo que estão associadas por exemplo: à transmissão do vírus da influenza [4], gastrites e cancro do estômago [5], síndrome da imunodeficiência (HIV) [6] entre outros, iii) reconhecimento imunológico de células tumorais e iv) reconhecimento de tecidos celulares de origem externa após transfusão de órgãos [7].

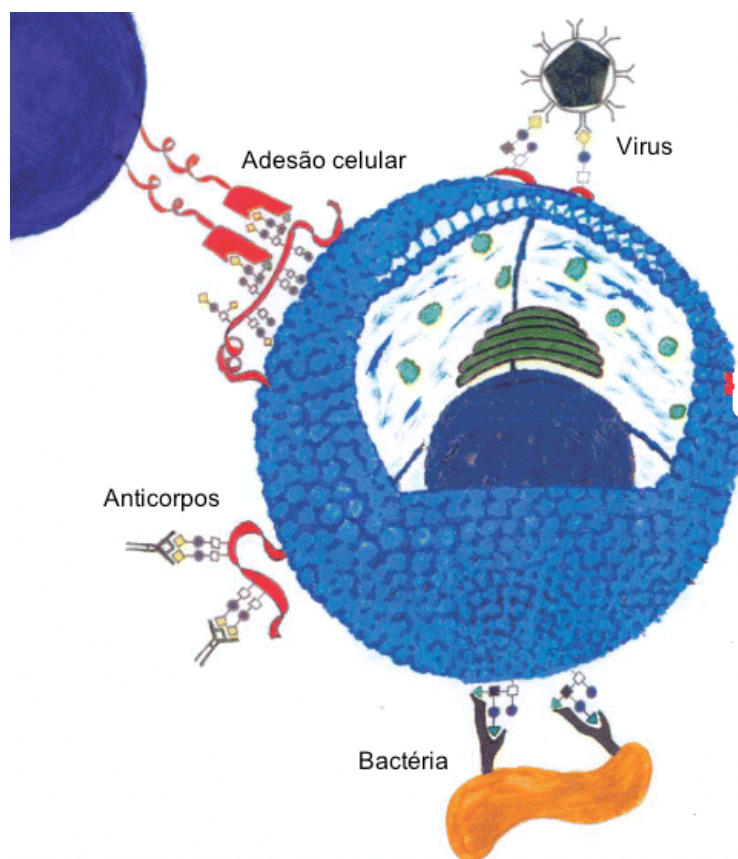


Figura I.1. Interação proteína-carboidrato. Esquema de diversos fenómenos biológicos comandados por interações proteína-carboidrato à superfície celular, tais como ligação célula-célula e célula-anticorpo, bem como a adesão de célula-microrganismo (bactéria ou vírus). Figura adaptada da referência [8].

Deste modo, para entender como estes processos ocorrem é de grande importância esclarecer os aspectos dos detalhes moleculares das interações proteína-carboidrato. Neste

trabalho, para se obter mais conhecimentos sobre estes processos, serão utilizados como modelo de estudo duas proteínas de *Bacillus subtilis* envolvidas no reconhecimento de carboidratos: uma proteína de ligação a carboidratos (AraN [9] – proteína de reconhecimento de substratos (SBP) que é componente de um sistema de transporte do tipo ABC) e uma fosfatase de carboidratos do tipo ADH (AraL [10]).

A bactéria Gram-positiva *Bacillus subtilis* encontra-se envolvida na secreção de um consórcio de enzimas grande parte delas hidrolases glicosídicas, que facilitam a despolimerização da parede celular vegetal em oligossacáridos mais simples. Este irão ser transportados para o interior celular, sendo subsequentemente utilizados como fonte de carbono para vias biossintéticas ou para fins energéticos através da fosforilação dos mesmos e sua incorporação na glicólise ou na via das pentoses fosfatadas [11].

I.1. L-Arabinose e a parede celular

A parede celular vegetal é uma entidade estrutural complexa, tendo como uma das suas principais funções conferir resistência mecânica ao stress osmótico. A mesma inclui três tipos de arquiteturas compostas por carboidratos: a celulose, hemicelulose e pectina [12].

Resíduos de L-arabinose podem ser encontrados na parede celular sob a forma de homopolissacáridos (arabinano) ou heteropolissacáridos (arabinoxilano e arabinogalactano), sendo estes alvos de análise dado serem substrato para os produtos codificados pelo genes pertencentes ao operão *araABDLMNPQ-abfA*, pese embora a *Bacillus subtilis* utilize um conjunto mais lato de substratos através da utilização de enzimas e transportadores codificadas noutros genes e operões (ver tabela 1 presente no artigo de revisão de Stülke *et al* [11]).

No que diz respeito aos arabinanos presentes na pectina, estes podem ser lineares ou ramificados, constituídos por subunidades de α -L-arabinofuranose polimerizadas por ligações glicosídicas (1 \rightarrow 5), sendo que as ramificações podem ocorrer nas posições O2 ou O3, como é possível constatar na figura I.2.

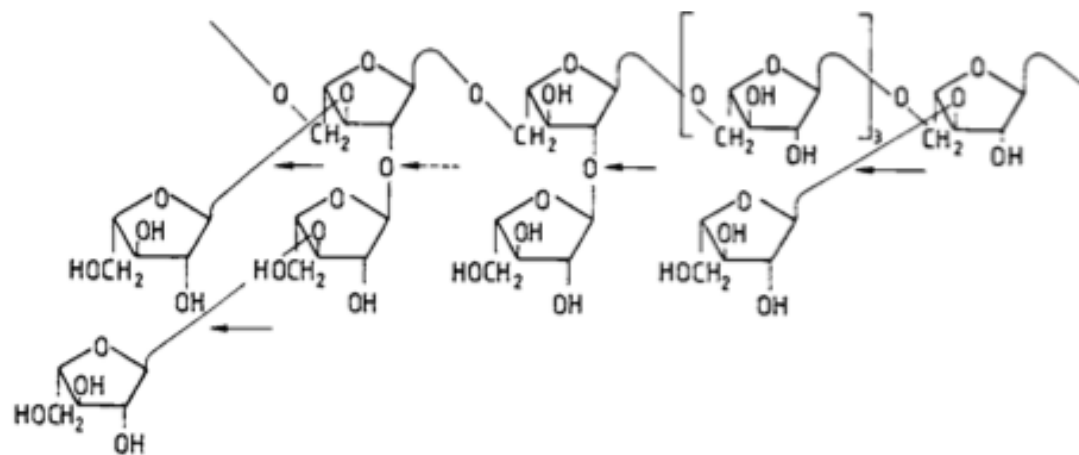


Figura I.2. Estrutura do arabinano. Esquema do arabinano ramificado. Figura retirada da referencia [13].

Outro açúcar presente na pectina, o arabinogalactano tem, tal como o arabinano, como principal função conferir flexibilidade à pectina. Estruturalmente, está subdividido em três classes,

sendo genericamente constituído por uma cadeia linear β -D-galactopiranosose polimerizada por ligações glicosídicas 1 \rightarrow 4. O arabinogalactano do tipo I apresenta ramificações de polímeros de α -1,5-L-arabinofuranose na posição O4, ao passo que do tipo II apresenta a mesma ramificação nas posições O3 e O6. Finalmente, o arabinoxilano pode ser encontrado em hemiceluloses. Este é constituído por β -D-xilanopiranosose ligado em 1 \rightarrow 4, sendo ramificado nas posições O2 ou O3 por homopolímeros de α -L-arabinofuranose [13].

Como foi dito anteriormente, a *Bacillus subtilis* produz um consorcio de arabinases com atividade endo- e exo-hidrolítica com o intuito de libertar oligómeros de arabinose e L-arabinose que serão subsequentemente transportadas para o interior da célula bacteriana e consequentemente metabolizadas. Especificamente, tanto a AbnA [14] como a Abn2 [15] pertencem à família 43 das glicosil hidrólases (GH43), e apresentam atividade endo-1,5- α -L-arabinanase. Estas enzimas extracelulares promovem a hidrólise das ligações glicosídicas levando à formação de arabinoligossacáridos. Estes últimos são transportados para o interior da célula pela ação de um transportador do tipo ABC (*ATP binding cassette transporters*) – AraNPQ-MsmX [9]. Por seu turno as enzimas citoplasmáticas α -L-arabinofuranosidases AbfA e Abf2 [16] pertencentes à família GH51, promovem a libertação de L-arabinose. Estas enzimas apresentam uma atividade colaborativa, isto porque a primeira tem uma maior atividade para ligações glicosídicas 1 \rightarrow 5 entre arabinofuranoses, ao passo que a segunda apresenta uma maior atividade para ligações 1 \rightarrow 2 e 1 \rightarrow 3. Na figura I.3, encontra-se resumido o processo até aqui descrito.

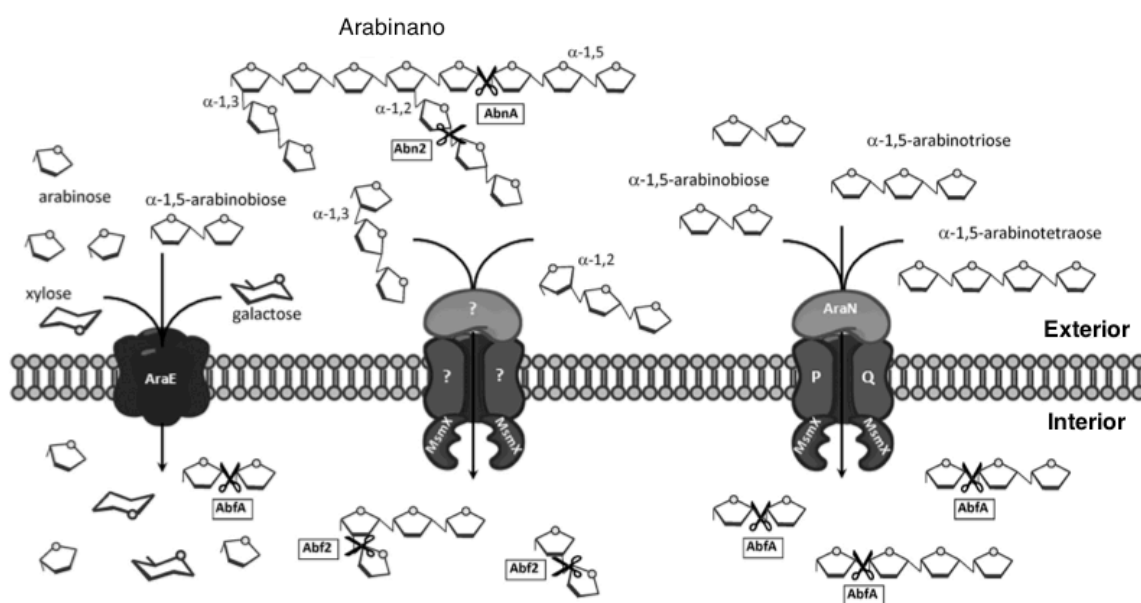


Figura I.3. Transporte primário e secundário de arabinoligossacáridos em *Bacillus subtilis*. Esquema ilustrativo dos processos de despolimerização do arabinano e consequente transporte de arabinoligossacáridos e arabinose através da ação de transportadores de membrana. Já no interior o arabinoligossacárido é continuamente despolimerizado a L-arabinose. Imagem adaptada da referência [9].

Finalmente a L-arabinose é metabolizada pela ação de enzimas codificadas no operão *araABDLMNPQ-abfA*, nomeadamente pelo produto dos genes *araA* (L-arabinose isomerase), *araB* (L-ribulocinase) e *araD* (L-ribulose-5-P-4-epimerase) [17].

I.2. Transporte de arabinoligossacáridos para o interior da célula

No que toca aos mecanismos adoptados no transporte através da membrana plasmática, este pode ser passivo (difusão simples ou facilitada), ou ativo (primário ou secundário), na tabela I.1 encontram-se resumidas as propriedades destes mesmos transportadores.

Tabela I.1. Topologia de transporte. Mecanismos de transporte e resumo de algumas características dos diferentes tipos de transportadores, bem como exemplos dos mesmos.

Propriedades	Difusão simples	Difusão facilitada	Transporte ativo primário	Transporte ativo secundário
Proteína transportadora	-	+	+	+
Transporte de soluto contra gradiente eletroquímico	-	-	+	+
Acoplado à hidrólise de ATP	-	-	+	-/+
Dirigido pelo cotransporte de substrato a favor de gradiente eletroquímico	-	-	-	+
Exemplo de moléculas transportadas	O ₂ ; N ₂ ; Glicerol	Glucose	Maltodextrina Ca ²⁺	H ⁺ /Lactose
Exemplo de transportadores	-	GLUT1[18]	MdxEFG (ABC)[19] SERCA (ATPase)[20]	LacY[21]

⁺Presença da propriedade

⁻Ausência da propriedade

O transporte passivo pode ocorrer mediante uma difusão simples, no qual o substrato irá fazer uso das suas características lipofílicas para se dissolver e atravessar a membrana, sendo que quanto maior a constante de partição mais rápida é a difusão. Como é fácil de prever há um conjunto de moléculas e iões que face à sua polaridade ou tamanho não conseguem permear a membrana sem o auxílio de proteínas transportadoras (difusão facilitada ou transporte ativo).

Um denominador comum à difusão simples e facilitada, é que em ambos os transportes os solutos movem-se a favor de um gradiente eletroquímico, tornando o processo termodinamicamente espontâneo. Relativamente ao transporte ativo este é feito contra um gradiente eletroquímico e requer energia celular. Este pode subdividir-se em transporte ativo primário, no qual a hidrólise de ATP irá energizar o processo, contribuindo para esta família de transportadores as ATPases do tipo P [20], bem como os transportadores do tipo ABC [19]. Relativamente ao transporte ativo secundário, o transporte de um substrato S₁ contra o seu gradiente eletroquímico é conseguido em função do mesmo estar acoplado ao transporte de um segundo substrato S₂ a favor de um gradiente de concentrações e do potencial de membrana ($\Delta G^{S_1} > 0$; $\Delta G^{S_2} < 0$). O transporte de ambos os substratos pode estar acoplado de forma a que ambos os substratos se movam na mesma direção (simporte) ou em direções opostas (antiporte) [22].

Muitas das proteínas embebidas na membrana plasmática, têm como principal função a atribuição de permeabilidade, formando transportadores e canais (transporte ativo e passivo) essenciais em diversos processos biológicos, de assimilação de nutrientes [9], no estabelecimento de potenciais elétricos [23] para a propagação dos mesmos ao nível dos neurónios, na excreção de xenobióticos, e na multirresistência a antibióticos [24]. Relativamente aos transportadores, estes encontram-se subdivididos em duas famílias: SLC e ABC [25, 26].

I.2.1. Transportadores do tipo ABC

Em *Bacillus subtilis*, os oligossacáridos resultantes da hidrólise prévia dos arabinanos são transportados através de um transportador do tipo ABC. Estes constituem uma superfamília de proteínas integrais de membrana, responsáveis pelo transporte de um largo espectro de substratos através da membrana plasmática. Para tal, estes transportadores acoplam a hidrólise de ATP à translocação de substratos seja este um processo de importação ou exportação, através de um mecanismo análogo [27, 28].

Até à data, no que diz respeito a importadores, estes apenas foram encontrados em procariontes sendo que estes necessitam de uma proteína responsável pela captura do substrato com grande afinidade e especificidade, e subsequente entrega aos componentes integrais de membrana. No caso do transportador AraNPQ-MsmX de *Bacillus subtilis* esta proteína é codificada pelo gene *araN*. Relativamente aos exportadores, estes encontram-se também em eucariontes [29].

Atualmente, existe um grande interesse em desvendar os detalhes do mecanismo de funcionamento destas proteínas membranares, visto que mutações das mesmas estão envolvidas numa vasta gama de doenças, bem como associadas à multirresistência a antibióticos[30-32].

I.2.1.1. Organização estrutural dos transportadores do tipo ABC

No que toca à arquitetura dos ABC's, estes são essencialmente compostos por quatro domínios, dois transmembranares (TMD) servindo estes de veículo transportador do substrato, e dois domínios de reconhecimento de nucleótidos (NBD) responsáveis pelo reconhecimento e hidrólise do ATP, como se encontra ilustrado na figura I.4.

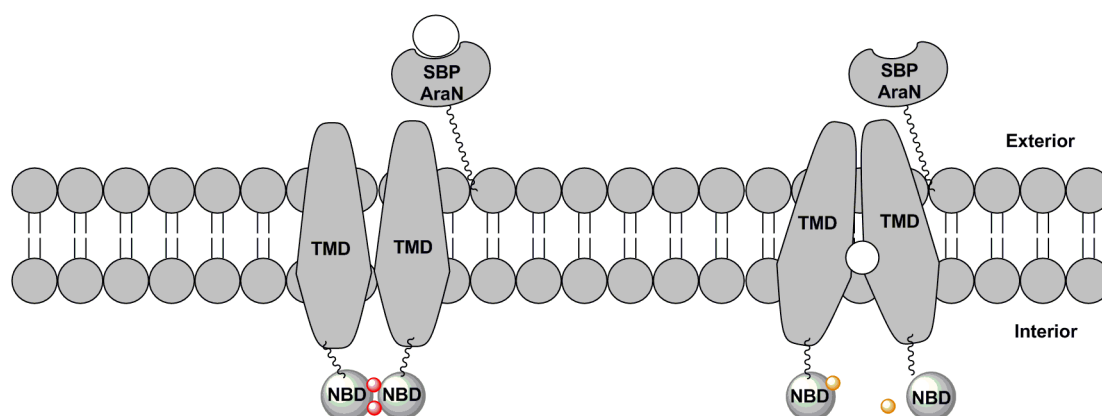


Figura I.4. Transportador do tipo ABC. Representação modular da organização dos transportadores ABC, composta por dois módulos TMD, dois módulos NBD e pelo menos uma SBP. Figura adaptada da referência [33].

Relativamente aos 4 domínios que compõem os transportadores ABC, estes podem apresentar diferentes arranjos consoante a sua função e consoante o tipo de organismo que o expressa. No caso dos transportadores ABC que funcionam enquanto importadores, os quatro domínios são expressos como cadeias polipeptídicas separadas. Contudo, os exportadores procariontes, expressam as subunidades TMD-NBD fundidas, sendo o transportador um homo ou heterodímero, já os eucariontes apresentam o transportador enquanto uma única unidade polipeptídica, com todos os domínios intervenientes unidos [34]. Dentro da categoria de importadores do tipo ABC existem duas subcategorias [35]:

1. Importadores do tipo I: Responsáveis pelo transporte de iões, carboidratos, aminoácidos.
2. Importadores do tipo II: Responsáveis pelo transporte de substratos de maiores dimensões comparativamente ao tipo I, entre os quais se incluem o transportador de vitamina B12 (BtuF) [36, 37].

I.3. Proteínas de ligação ao substrato- SBP

As SPB, têm como principal função o auxílio à captação de nutrientes através da membrana citoplasmática. Tal como é ilustrado por *Shuman et al* [38], a deleção destas proteínas previne o transporte de nutrientes através da membrana citoplasmática mesmo na presença das restantes proteínas que constituem o transportador do tipo ABC. Neste estudo, uma estirpe *wild-type* de *Escherichia coli* apresenta um tempo de duplicação de 60 min num meio 5.5mM em maltose (substrato captado pelo transportador). A deleção da SBP provoca neste organismo ausência de crescimento em 5.5mM e 25mM de maltose [38].

No entanto, os locais de reconhecimento não se limitam apenas às SBP's. Para além destas, as subunidades transmembranares apresentam também elas locais de reconhecimento de nutrientes. Para evidenciar este facto os autores procederam à mutação de subunidades transmembranares do transportador acima referido, sendo que estes mutantes evidenciaram a capacidade de transportar nutrientes na ausência de SBP's. A mutação das subunidades transmembranares levou à exposição dos locais de reconhecimento nas mesmas, algo que na estirpe *wild-type* apenas acontece fruto do mecanismo de reconhecimento das SBP's na forma carregada que através da interação proteína-proteína promove a abertura dos TMD's [39].

Historicamente, estas proteínas começaram por ser classificadas como proteínas resistentes ou sensíveis a um choque osmótico [40]. Assim, as SBP's presentes em bactérias Gram-negativas irão encontrar-se solúveis no espaço periplasmático, perdendo o transportador a sua capacidade de translocação de substratos após o choque osmótico, estas SBP's dizem-se sensíveis ao choque osmótico. Por outro lado as bactérias Gram-positivas irão apresentar SBP's ancoradas à membrana citoplasmática via domínio N-terminal face à sua natureza lipoproteica [41, 42], ou poderão encontrar-se ancoradas diretamente ao N ou C-terminal da região transmembranar do transportador [43].

No que diz respeito às SBP's presentes em bactérias Gram-positivas ancoradas às regiões N ou C-terminais, estas podem ainda estar sob a forma de homodímero, sendo que por cada TMD's

será possível encontrar dois SBP's. Cada transportador é constituído por duas subunidades transmembranares, podendo então estes sistemas de transporte apresentar 4 locais de ligação nas proximidades dos transportadores [44]. A presença destes múltiplos locais de ligação de substratos podem trazer vantagens aos sistemas de transporte sejam devido à ampliação da especificidade do transportador, ou devido ao aumento da capacidade na translocação de substratos. No primeiro caso, para o mesmo sistema de transporte cada SBP é responsável pelo reconhecimento de um substrato diferente providenciando assim ao sistema uma reconhecimento mais abrangente, como é evidenciado na secção seguinte. No segundo caso, os SBP's poderiam interagir alternadamente com o transportador, promovendo um aumento da taxa de transporte face ao aumento do numero de SBP, as unidades responsáveis pelo reconhecimento dos substratos a ser transportados.

No que toca aos substratos translocados, estas SBP's apresentam um grande afinidade por uma ampla gama substratos, como são o caso de mono e oligossacáridos [45], aminoácidos [46] e pequenos péptidos [47, 48], vitaminas [37], metais, aniões [49], catiões [50]. Os SBP's podem ainda ser específicos para um família de substratos estruturalmente semelhantes, ou apresentar um ampla gama de substratos com diferenças estruturais como é o caso do sistema de transporte MalEFG-MsmK de *Thermus thermophilus*. Este apresenta uma SBP com afinidade para diferentes carboidratos (maltose, trehalose, sucrose e palatinose) [51]. No que toca à versatilidade de substratos translocados esta pode também ser conseguida através de um transportador que reconheça múltiplos SBP's, estes últimos com especificidades para diferentes substratos, como é o caso do sistema de transporte Ami presente na bactéria *Streptococcus pneumoniae*, responsável pelo transporte de oligopeptídeos, através de três SBP's sendo estas: AmiA, AliA e AliB que apresentam seletividades diferentes para diferentes oligopeptídeos [52]. Os SBP's com maior poder discriminatório de substratos são aqueles responsáveis pela ligação de oxianíons, como é o caso das proteínas específicas de ligação a sulfato (*SulphateBP*) e de ligação a fosfato (*PhosphateBP*).

I.3.1. AraN- Proteína de reconhecimento de arabinoligossacáridos

A proteína AraN é uma SBP que em conjunto com a AraPQ, bem como com a proteína com atividade de ATPase MsmX, formam um transportador do tipo ABC. As proteínas AraPQ irão compor os TMD's ao passo que os NBD's serão constituídos pela MsmX. Este transportador de membrana, encontra-se envolvido na translocação de α -1,5-arabinoligossacáridos, como é possível observar na figura I.3.

Tal como o transportador de maltose/maltodextrina MalEFGK₂, o transportador AraNPQ-MsmX não tem capacidade de reconhecimento de unidades monoméricas do substrato translocado. Em suma, este transportador é responsável pela translocação de unidades de L-arabinose até quatro subunidades, sendo que a α -1,5-arabinotriose e a α -1,5-arabinotetraose, são transportados única e exclusivamente por este transportador. Cabe ao transportador ativo primário AraNPQ-MsmX e ao transportador secundário simporte de hidrogénio (AraE) a translocação de arabinobiose. Na tabela I.2, encontram-se evidenciadas as cinéticas de crescimento de diferentes estirpes nestes substratos [9].

Tabela I.2. Taxas de crescimento de diferentes estirpes de *B. subtilis*. Efeito das diferentes mutações, na presença de diferentes fontes de carbono. NG, Sem crescimento. – Não determinado.[9]

Fonte de carbono	168T ⁺ (WT)	IQB496 (Δ araN)	IQB611 (Δ araNPQ)	IQB608 (araE::kan)	IQB610 (Δ araNaraE::kan)
Arabinose	67,0 \pm 0,2	66,6 \pm 0,9	63,7 \pm 2,9	NG	NG
Arabinobiose	112,4 \pm 10,8	263,8 \pm 37,2	258,2 \pm 46,1	178,6 \pm 9,9	NG
Arabinotriose	98,2 \pm 10,0	NG	NG	-	-
Arabinotetraose	92,4 \pm 4,5	NG	NG	-	-

I.3.2. Topologia estrutural de SBP

As SBP's apesar de não apresentarem uma extensa homologia de sequencia primária de aminoácidos, apresentam todas elas entre si um grande homologia ao nível a estrutura terciária, sendo estas constituídas por dois lobos globulares N e C, presentes respectivamente nas regiões N e C-terminais da proteína [53]. Estes dois lobos encontram-se ligados por segmentos peptídicos, que compõem o local de ligação ao substrato. A ligação do substrato à proteína é acompanhado por um movimento de torção/rotação que irá coincidir com o fecho dos dois domínios sobre o substrato, como é observado na figura I.5.

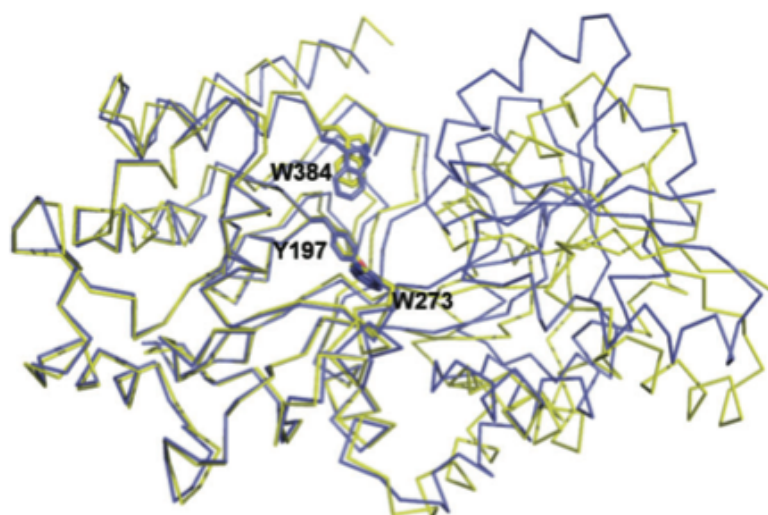


Figura I.5. Representação da SBP Malx de *Streptococcus pneumoniae*. A azul está a *backbone* Ca após reconhecimento de maltoheptaose. A amarelo está representada a proteína sem ligando. Há uma aproximação do domínio N-terminal após reconhecimento do ligando no local de ligação conservado (W384, Y197, W273). Figura retirada da referência [54].

Estruturalmente estes lobos apresentam folhas- β no núcleo, sendo estas flanqueadas em ambos os planos da folha por hélices- α num arranjo *Rosemann fold*. Como foi dito anteriormente, estes dois lobos encontram-se ligados por pequenos segmentos peptídicos. Com base no número de segmentos peptídicos surge uma classificação de SBP's. O grupo I é constituído por SBP's cujos lobos se encontrem ligados por três pequenos segmentos peptídicos. O grupo II apresenta dois segmentos peptídicos a conectar ambos os domínios, e finalmente existe o grupo III cujos domínios N e C-terminais encontram-se ligados por um único segmento peptídico [55, 56]. Na tabela I.3, encontra-se resumida esta classificação de SBP's tendo em conta o tipo de substrato transportado ao longo da membrana citoplasmática [56].

Tabela I.3. Classificação das SBP's. Categorização consoante o grupo de proteínas a que pertence, bem como o substrato translocado.

SBP	Grupo	Substrato	Origem
AraF	I	L-Arabinose	<i>Escherichia coli</i>
RbsB	I	D-Ribose	<i>Escherichia coli</i>
MglB	I	Glucose/ galactose	<i>Salmonella typhimurium</i>
LivJ	I	Leucina/ isoleucina/ valina	<i>Escherichia coli</i>
LivK	I	Leucina	<i>Escherichia coli</i>
HisJ	II	Histidina	<i>Escherichia coli</i>
GlnH	II	Glutamina	<i>Escherichia coli</i>
ArgT	II	Lisina/ arginina/ ornitina	<i>Salmonella typhimurium</i>
MalE	II	Maltose/ maltodextrina	<i>Escherichia coli</i>
Sbp	II	Sulfato	<i>Salmonella typhimurium</i>
PstS	II	Fosfato	<i>Escherichia coli</i>
BtuF	III	Vitamina B ₁₂	<i>Escherichia coli</i>
PsaA	III	Mn ²⁺	<i>Streptococcus pneumonia</i>
TroA	III	Zn ²⁺	<i>Treponema pallidum</i>

I.3.3. Alteração conformacional das SBP's e translocação de substrato

O reconhecimento de substrato por parte das SBP's que ocorre nas regiões de ligação entre o domínio N e C-terminal, é acompanhado por uma torção/rotação (*hinge bending motion*) de um conjunto de resíduos que se encontram nesta região. Este mecanismo será responsável pela rotação dos dois domínios anteriormente referidos, e consequente fecho sobre o substrato- *Venus flytrap mechanism* [57]. Este fecho está em última análise associado a modificações conformacionais da proteína SBP, tornando-a capaz de interagir subsequentemente com as subunidades transmembranares do transportador ABC, permitindo a abertura das mesmas e exposição de um segundo local de ligação ao substrato nas subunidades transmembranares.

Para perceber quais os resíduos que maior contribuição têm na abertura e consequente processo de reconhecimento do substrato a ser translocado, foram feitas medições das perturbações dos ângulos de torção da cadeia principal ao nível do C_α. Para a torção desta ligação após abertura ou fecho da proteína, irão contribuir as rotações ao nível da ligação química C_α e o carbono carbonilo (-CO), sendo estas rotações governadas pelos ângulos diedros Ψ (psi). Bem como a rotação ao nível da ligação química C_α -NH, rotações estas governadas pelo ângulo Φ (phi). A abertura desta proteína irá culminar com uma perturbação dos ângulos de torção que decoram as fronteiras entre os dois domínios, e que em última análise são constituídas pelo segmentos peptídicos que os conectam [58].

Como foi referido anteriormente, um ponto crítico do transporte mediado por ABC's com auxílio de SBP's reside na capacidade das subunidades transmembranares do ABC distinguirem a conformação aberta (vazia) e fechada (com ligando) da SBP. A ligação do substrato ao local de

ligação irá induzir o fecho das SBP's. Esta alteração conformacional irá providenciar a exposição dos locais de reconhecimento proteína-proteína (BPD-TMD). Assim, o mecanismo orquestrado de reconhecimento do substrato pela SBP, irá ser essencial para a discriminação de SBP fechada e carregada com substrato, dos demais SBP's livres [58].

A importância destas modificações conformacionais ao nível do reconhecimento do substrato e consequente translocação ficou evidente em estudos de mutagénese da MBP (*Maltose Binding Protein*) de *Escherichia coli* [58], pela substituição dos resíduos Leu107 e Ile111, por aminoácidos de maiores dimensões. Estes resíduos encontram-se próximo das regiões que irão sofrer a rotação/torção. Na forma aberta estes resíduos aproximam-se e na forma fechada afastam-se. A substituição destes resíduos por aminoácidos com cadeias laterais de maiores dimensões irá promover o deslocamento do equilíbrio das formas abertas para as fechadas. Presume-se que a SBP na ausência de ligando, se encontre num equilíbrio entre a forma aberta e a forma fechada, e que a ligação do substrato desloque este equilíbrio no sentido da conformação fechada, aquela reconhecida pelo transportador transmembranar. Em conclusão as modificações conformacionais ao nível dos segmentos peptídicos que se encontram a ligar ambos os domínios irão ter implicações não apenas no reconhecimento molecular do substrato, mas também na sua libertação aquando do transporte. A associação do SBP às subunidades transmembranares do transportador irá levar à hidrólise de ATP das subunidades NBD's, possibilitando a abertura das SBP's com libertação consequente do substrato para as TMD's [59]. Todo o mecanismo de transporte do substrato é em última análise orquestrado pelo reconhecimento molecular proteína-ligando.

I.3.4. Especificidade das SBP's

No que toca à especificidade de substratos aos quais as SBP's se conseguem ligar, esta é conseguida primariamente através de ligações por pontes de hidrogénio.

Relativamente a carboidratos, os grupos funcionais hidroxilo (-OH) estão expostos e disponíveis para interagir com a proteína, nomeadamente com resíduos cujas cadeias laterais polares têm a capacidade de funcionar enquanto dadores ou aceitadores de pontes de hidrogénio. Dentro das pontes de hidrogénio, observaram-se também a formação das mesmas de um modo cooperativo, no qual cada hidroxilo do açúcar se encontra a estabelecer uma ponte de hidrogénio com uma amida (-NH) –dador de pontes de hidrogénio- bem como com um carbonilo (-CO)- aceitador de pontes de hidrogénio. Para além destas também é recorrente a formação de pontes de hidrogénio bidentadas, que acontece quando dois hidroxilos consecutivos no açúcar estabelecem pontes de hidrogénio com ácidos aspártico ou glutâmicos. A formação de extensas redes de pontes de hidrogénio é também mediada por moléculas de água que fazem a “ponte” entre resíduos polares do ponto de vista proteína e hidroxilo presentes no carboidrato [60, 61]. Os eventos de reconhecimento molecular são resultado da combinação de resíduos da proteína e do ligando. Nesta medida os resíduos passíveis de se encontrarem a estabelecer pontes hidrogénio com o carboidrato serão de natureza polar tais como a o ácido glutâmico e ácido aspártico bem como a asparagina, glutamina, arginina e histidina.

A interação entre proteínas e carboidratos também ocorre através de interações hidrofóbicas. A disposição de grupos -CH do açúcar pode levar à formação de regiões hidrofóbicas, que irão interagir com resíduos cuja cadeia lateral é também ela hidrofóbica (triptofano, tirosina e

fenilalanina) através de interações CH- π *stacking*. Estas ocorrem através das cadeias laterais aromáticas e os átomos de hidrogénio do anel que compõe o açúcar [55].

Para além das anteriores, os contactos dipolares de van der Waals também desempenham um papel fundamental no reconhecimento de substratos.

No que toca ao reconhecimento de aminoácidos este é efectuado tendo por base a seletividade para com a cadeia lateral específica de cada aminoácido. O transportador de leucina/isoleucina/valina (LIVPB), consegue efetuar o reconhecimento destes aminoácidos face à presença de um local de ligação decorado de aminoácidos também eles tais como os primeiros cujas cadeias laterais são de natureza hidrofóbica. Ainda no que ao reconhecimento de aminoácidos diz respeito, o reconhecimento da histidina por parte do seu transportador HBP, é conseguida por um *stacking* do grupo imidazole com uma tirosina, bem como pela formação de pontes de hidrogénio entre os dois azotos desse mesmo grupo imidazole com aminoácidos polares que residem no local de ligação [62]. O reconhecimento de aminoácidos lisina/ornitina/arginina (LAOBP) é efectuado maioritariamente através da formação de interações electroestáticas entre os resíduos carregados na cadeia lateral, formando assim pontes salinas [63]. A seletividade observada para as proteínas evidenciadas acima consoante a natureza da cadeia lateral não se verificam no entanto para transportadores de di (DppA) e oligopeptidos (OppA), que têm um reconhecimento mais lato de substratos translocados.

As proteínas de ligação a sulfatos e fosfatos são um caso de especificidade único. Estas proteínas conseguem fazer a distinção entre oxianíons estruturalmente muito semelhantes como é o caso do sulfato e fosfato, sendo que esta distinção está ultimamente associada ao estado de protonação dos mesmos, sendo que o fosfato se encontra protonado até um pH=13 ao passo que o sulfato encontra-se ionizado a pH=3.

Ao contrário dos casos anteriores de interação proteína-carbohidrato e proteína-proteína, no qual as interações por parte do receptor são efectuadas maioritariamente através das cadeias laterais dos resíduos que compõem o local de ligação, no caso das proteínas de reconhecimento de sulfato e fosfato o reconhecimento molecular é conseguido através de pontes de hidrogénio entre os oxigénios do anião e as amidas e carboxilos polarizados ao nível da cadeia principal da proteína.

Em última instância, os grupos hidroxilo das cadeias laterais também podem funcionar como dadores de pontes de hidrogénio, no entanto o reconhecimento cabe praticamente inteiramente aos –NH da cadeia principal. Um ponto essencial no processo de reconhecimento do fosfato comparativamente ao sulfato como foi referido anteriormente reside no facto de este apresentar um oxigénio protonado que irá funcionar como dador de pontes de hidrogénio com um grupo carboxilo presente no local de ligação de proteínas de ligação a fosfatos. No total, os complexos de reconhecimento de fosfatos por SBP's envolvem a formação de 12 pontes de hidrogénio.

Ao contrário do fosfato, o sulfato a pH neutro encontra-se completamente ionizado, sendo então que a SBP de ligação a este sulfato não irá requerer o aceitador de pontes de hidrogénio para ligar o hidroxilo –OH como se encontra presente no fosfato a pH neutro. Assim o reconhecimento do sulfato requer a formação de 7 pontes de hidrogénio, sem que a proteína SBP específica deste

oxianião contenha resíduos carregados negativamente a “decorar” o local de ligação, como tem obrigatoriamente de acontecer no caso do fosfato.

I.4. Desalogenase de haloalcanos

A transferência de grupos fosfatos, é um processo proeminente na natureza, observado nos mais diversos fenómenos biológicos. Um dos mecanismos mais comuns na regulação da atividade de proteínas é efectuado através da fosforilação de grupos hidroxilo de aminoácidos (alterando com isso a carga da proteína ou conformação do local de ligação, impedindo o reconhecimento do ligando) tais como serinas, treoninas e tirosinas, sendo que as atividades antiparalelas de cinases e fosfatases irão providenciar mecanismos de ativação ou desativação das mesmas. Eventos celulares tais como metabolismo primário, transcrição, ciclo-celular, apoptose, bem como funções neuronais e imunológicas, são alvo de regulação por parte de cinases e fosfatases [64]. Na tabela I.4, encontram-se catalogadas as enzimas presentes na família de DHA's.

Tabela I.4. Topologia de enzimas DHA. Enzimas presentes na família DHA, e o substrato clivado pelas respectivas.

Tipo de enzima	Substrato clivado
Desalogenase ácida de halogéneos	C-Cl
Fosfonatase	C-P
Fosfatase	CO-P
Fosfomutase	CO-P
ATPase	PO-P

As hidrólases de esteres de fosfato (fosfatases), apresentam uma topologia *Rossmann fold* α/β , com um conjunto de motivos estruturais conservados envolvidos na ligação ao cofator Mg(II), ao reconhecimento do substrato, bem como à catálise ácido-base e ataque nucleofílico ao organofosfato.

Na tabela I.5 é possível analisar as diferentes funções executas por resíduos conservados nos diversos domínios das DHA's.

Tabela I.5. Motivos estruturais envolvidos na hidrólise do éster de fosfato. Resíduos conservados nos diferentes motivos, bem como a função que desempenham.

Domínio	Resíduos	Função
Domínio I	Asp	Catálise
Domínio II	Ser/Thr	Ligação ao fosfato
Domínio III	Lys/Arg	Ligação ao fosfato
Domínio IV	Asp/Glu	Coordenação ao Mg(II)

Relativamente ao mecanismo de hidrólise este é conseguido mediante um conjunto de etapas, evidenciadas na figura I.6. Numa primeira fase há uma transferência do grupo fosforilo para a cadeia lateral carboxílica de um ácido aspártico (nucleófilo). Numa segunda fase o grupo fosforil é

transferido para uma molécula de água. Para além do ácido aspártico (n), existe um segundo ácido aspártico (n+2) que desempenha um papel de ácido/base. Numa primeira fase funciona como ácido ao protonar o grupo de saída (-OR) e posteriormente como base para desprotonar o segundo nucleófilo (água). Estes dois resíduos, envolvidos no processo de catálise encontram-se posicionados no domínio I [65].

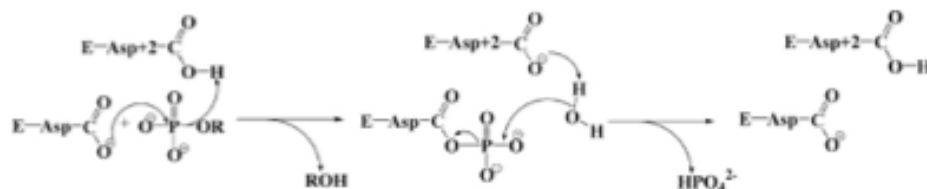


Figura I.6. Mecanismo de hidrólise de fosfatos. Diferentes etapas da hidrólise do fosfato (CO-P) promovida pela acção da DHA. Figura retirada da referencia [66].

I.4.1 AraL- Fosfatase de carboidratos fosforilados

A AraL é uma fosfatase que apresenta atividade para um conjunto de carboidratos fosforilados (*D*-xylulose-5-fosfato, glucose-6-fosfato, *D*-arabinose-5-fosfato, frutose-6-fosfato, frutose-1,6-bifosfato, galactose-1-fosfato e para-nitrofenolfosfato), evidenciando atividade máxima a pH=7, à temperatura de 65°C na presença de obrigatória de MgCl₂ (15mM), desempenhando este um papel funcional na atividade catalítica da enzima [10].

No estudo efectuado por *Godinho et al* [10], verificou-se através de ensaios de mutagenese, que o gene que codifica para a proteína AraL, não se encontra envolvido na utilização de *L*-arabinose. Face ao facto do gene *araL* se encontra presente no operão metabólico da arabinose *araABDLMNPQ-abfA*, bem como apresentar atividade para com *D*-xylulose-5-fosfato, resultante do catabolismo da *L*-arabinose (nomeadamente pela ação das enzimas AraABD), pode ser indicativo da utilização da AraL ao nível do metabolismo de carboidratos. Foi proposto para a AraL uma função putativa de destoxificação de intermediários fosforilados que se podem acumular ao nível das vias metabólicas (glicólise e via das pentoses fosfatadas), em situações particulares [10].

I.5. A Ressonância Magnética Nuclear no estudo da interação proteína-ligando

O reconhecimento molecular exercido entre a díade proteína-ligando, pode ser observado numa diversidade de processos biológicos que ocorrem na natureza. É essencial portanto desenvolver metodologias robustas para a compreensão qualitativa e quantitativa destes processos. Nesta medida a técnica de RMN apresenta vantagens comparativamente às demais possibilitando: i) determinação de *K_a* enquanto medida de afinidade da interação do complexo proteína-ligando, ii) determinação de epítomos de ligação quer do ponto de vista do ligando quer do ponto de vista da proteína, iii) caracterização estrutural de movimentos moleculares associados ao reconhecimento de substratos, processo recorrente ao nível das SBP's e em particular com AraN, iv) caracterização de processos de reconhecimento molecular dinâmicos (*turnover*) em catálise enzimática como é o caso da AraL.

O RMN apresenta como principais vantagens o facto de não ser uma técnica destrutiva (que é favorável quando a síntese de compostos é dispendiosa em termos económicos e em horas laborais, como é o caso da obtenção de algumas proteínas e em particular anticorpos). Em particular, o RMN é a técnica que maior semelhança permite atribuir ao sistema em estudo comparativamente às condições fisiológicas nas quais se encontram a proteína e o ligando, nomeadamente através do RMN em solução ou RMN *in vivo*. Finalmente a versatilidade da técnica fica também patente na capacidade que permite ao utilizador estudar sistemas de elevada afinidade entre o receptor e o ligando (estudos do ponto de vista da proteína) ou de média-baixa afinidade (estudos na ótica do ligando).

Conceptualmente, foram postuladas duas aproximações com o objectivo de estudar processos de reconhecimento molecular entre receptores biológicos e pequenas moléculas (ligandos), sendo que estas podem ser efectuadas i) sobre o ponto de vista do receptor molecular ii) do ponto de vista do ligando [67].

No que diz respeito aos estudos que pretendem elucidar quais as regiões do receptor que se encontram a interagir com o ligando, o parâmetro que usualmente é utilizado é o do desvio químico. O desvio químico apresentado por um núcleo é sensível ao ambiente químico envolvente, nomeadamente este irá ser o resultado do campo magnético aplicado, bem como dos pequenos campos magnéticos inferidos por parte do ligando. Assim, quando na presença do ligando, os núcleos presentes no local de ligação irão sofrer um efeito de blindagem ou desblindagem inerentes à aproximação do primeiro ao receptor (para maiores detalhes consultar secção VI.1.1. *Desvio químico*).

Tal como se encontra evidenciado na figura I.7., a análise das perturbações dos desvios químicos do ponto de vista da proteína [68, 69] requer duas premissas i) a atribuição inequívoca das ressonâncias espectrais, ii) estrutura tri-dimensional da proteína [70, 71]. A análise das perturbações dos desvios químicos pode ser acompanhada ao longo de uma titulação através da aquisição sucessiva de espectros ^1H - ^{15}N -HSQC garantida a previa marcação isotópica de todas as amidas presentes na proteína [72-76]. Esta última informação dá ao operador a informação qualitativa de que está a ocorrer reconhecimento molecular por parte da proteína, bem como quais os resíduos envolvidos nesse mesmo processo, e por outro lado informação quantitativa sobre a forma de uma constante de afinidade acerca do processo de reconhecimento, visto que as perturbações dos desvios químicos seguem uma cinética descrito pela equação de Michaelis-Menten com o aumento da concentração de substrato ao longo da titulação. Para além disto, os resíduos perturbados podem ser mapeados ao nível da estrutura tridimensional para identificar locais de ligação [77].

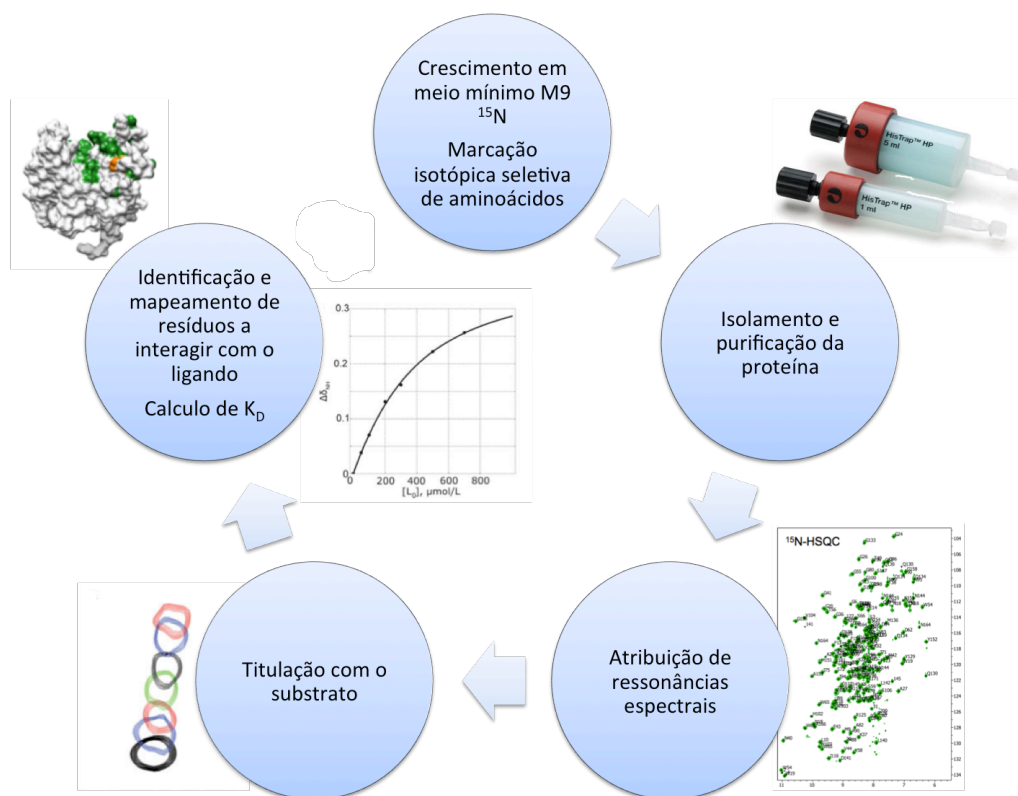


Figura I.7. Interação do ponto de vista da proteína. Estratégias utilizadas no estudo de interações sobre o ponto de vista da proteína. Numa etapa inicial são empregues diferentes estratégias de marcação isotópica (Meio mínimo M9 ou incorporação de aminoácidos marcados). Seguidamente a proteína de interesse é isolada e purificada. Consequentemente, após atribuição de todas as ressonâncias espectrais e cálculo de estrutura, é efectuada a titulação com um substrato de interesse, para se aferir a região de reconhecimento molecular e a constante de afinidade deste mesmo processo.

Alternativamente à marcação isotópica da proteína, caso esta apresente um número elevado de resíduos 50-100kDa, a estratégia de marcação isotópica tem de ser alterada, pois pese embora apenas apareçam aproximadamente o mesmo número de amidas face ao número de aminoácidos, a probabilidade de ocorrer sobreposição de ressonâncias espectrais é elevada. Assim uma das alternativas passa pela incorporação de aminoácidos previamente marcados pela estirpe de sobreexpressão [78, 79]. Os aminoácidos escolhidos devem ser aqueles que apresentem uma maior probabilidade de decorar a cavidade de ligação. Por outro lado a técnica de SEA-TROSY [80], pode também ser útil na medida que o espectro originado por esta técnica apenas irá apresentar as ressonâncias espectrais expostas ao solvente, eliminando assim resíduos mais enterrados no interior da proteína com funções estruturais no estabelecimento da estrutura terciária e que não têm funções no que ao reconhecimento molecular de substratos diz respeito. Assim esta técnica permite uma desconvolução espectral, permitindo uma análise inequívoca das perturbações dos desvios químicos, como é patente na figura I.8.

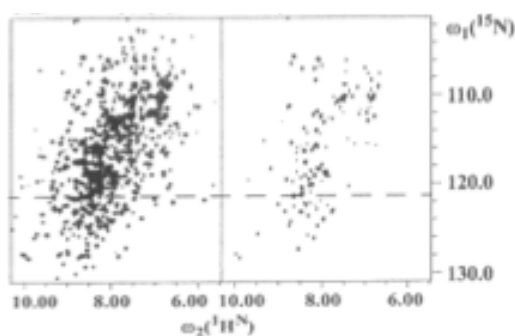


Figura 1.8. Espectro SEA-TROSY. Na região à esquerda da é possível observar o espectro TROSY, e à esquerda o espectro desconvoluído SEA-TROSY. Ilustração retirada da referencia [80].

A utilização de sondas paramagnéticas como é o caso do TEMPO (2,3,4,6-tetrametil-piperidina-1-oxil) [81] pode também ter um papel relevante na identificação de locais de ligação, face à diminuição da relaxação transversal T_2 aquando da interação dos spins nucleares nas vizinhanças de um eletrões desemparelhados [82]. Este técnicas pode ser de extrema relevância nomeadamente na elucidação do movimento dinâmico de *hinge-bending* descrito anteriormente, através da ligação de uma sonda paramagnética a um dos lobos que compõem a AraN.

No que diz respeito ao estudo da interação proteína-ligando do ponto de vista do ligando, esta pode ser efectuada essencialmente segundo duas abordagens. Numa primeira abordagem e em consonância com o que foi dito anteriormente, a interação pode ser estudado através da monitorização dos diferentes parâmetro do RMN tais como: desvio químico [83, 84], relaxação longitudinal T_1 , relaxação transversal T_2 (através da largura de linha) e difusão. Recentemente e na base do efeito de NOE surgiram novas abordagens baseadas na observação das alterações dos movimentos translacionais e rotacionais do ligando após reconhecimento por parte da proteína tais como o STD-RMN e o trNOE [85, 86]. A transferência de magnetização é considerado o *state-of-the-art* no que diz respeito ao estudo de interação sobre o ponto de vista do ligando quer através de STD-NMR ou WaterLOGSY. A principal vantagem deste estudo prende-se com o facto de apenas serem necessárias 10 μ g de proteína não marcada por ensaio.

Tal como foi evidenciado no ponto de vista da proteína, no ótica do ligando, as interações podem também elas ser detectadas através das perturbações do desvio químico, pela análise proposta anteriormente.

Em larga escala porém, a principal característica utilizada no que diz respeito ao estudo de interações do ponto de vista do ligando prende-se com as diferentes taxas de translação molecular em solução evidenciadas entre a proteína (alto peso molecular) e o ligando (baixo peso molecular). A figura 1.9. evidencia o impacto que o tamanho molecular e consequentemente o movimento molecular irão ter ao nível do NOE e da difusão enquanto parâmetros de caracterização de reconhecimento molecular [87].

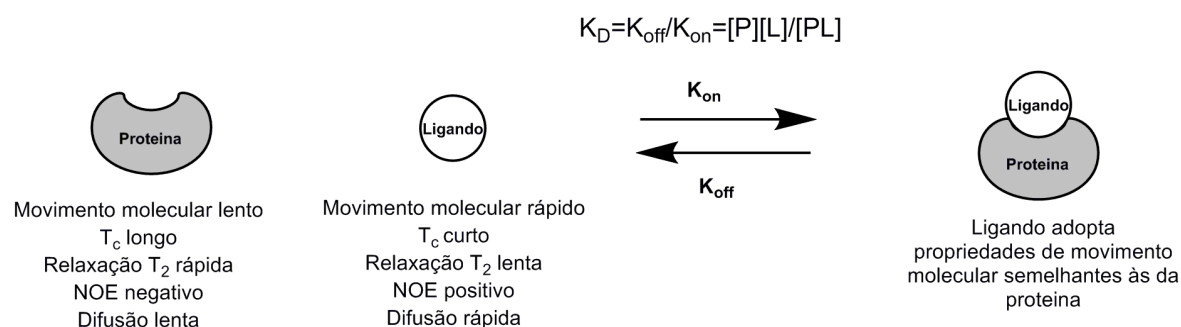


Figura I.9. Movimento translacional de diferentes moléculas em solução. Relação entre o tamanho da molécula em estudo e as suas características translacionais, rotacionais e difusionais em solução.

Relativamente a T_2 a largura de linha a meia altura de uma ressonância espectral de um ligando é bastante mais estreita que a largura de linha de uma estrutura macromolecular como é o caso da proteína. Após o reconhecimento do ligando por parte da proteína, este irá adoptar as características do segundo evidenciando proporcionalmente um aumento da largura de linha [87]. Por outro lado, efeitos de T_2 estão envolvidos na modelação da relaxação longitudinal T_1 , assim a relaxação T_1 de moléculas pequenas é também ela influenciada pela sua ligação a receptor macromoleculares, nomeadamente através da inversão do sinal de NOE (consultar secção VI.1.2 *Relaxação longitudinal T_1 e relaxação transversal T_2* , secção VI.1.3. *Efeito nuclear de Overhauser* e secção VI.1.4. *Movimento molecular e relaxação de spin*). Análogamente ao estudo das perturbações dos desvios químicos é possível titular o ligando com quantidade crescentes de proteína e aferir quais as regiões do ligando que se encontram a interagir com a proteína, bem como calcular o K_a desse processo, através da análise do alargamento de linha a meia altura [87].

Os ligandos de baixo peso molecular apresentam NOE's ligeiramente positivos ou nulos, no entanto após ligação ao receptor de molecular estes irão adoptar o seu movimento molecular e consequentemente ficaram com NOE's negativos. Assim, através da aquisição de uma espectro ^1H - ^1H -NOESY do ligando na ausência de proteína e de um segundo espectro na presença de proteína, a alteração de fase está associada ao reconhecimento molecular do ligando, sendo que as correlações cruzadas do espectro de NOESY permitirão ainda tirar conclusões acerca de alterações conformacionais passíveis de terem ocorrido do ponto de vista do ligando após terem entrado no local de ligação da proteína [86], como é possível observar na figura I.10.

Outras duas técnicas utilizadas no estudo de interações do ponto de vista do ligando são o STD-NMR [88, 89] e o WaterLOGSY [90, 91] que fazem uso da saturação seletiva da proteína em ambos os casos (no segundo esta saturação é mediada pela saturação da água) para identificar regiões do ligando que se encontrem a interagir com a proteína, pelo desaparecimento das restantes ressonâncias (STD-NMR), ou na alteração de fase das ressonâncias que não se encontram a interagir com a proteína no caso do WaterLOGSY.

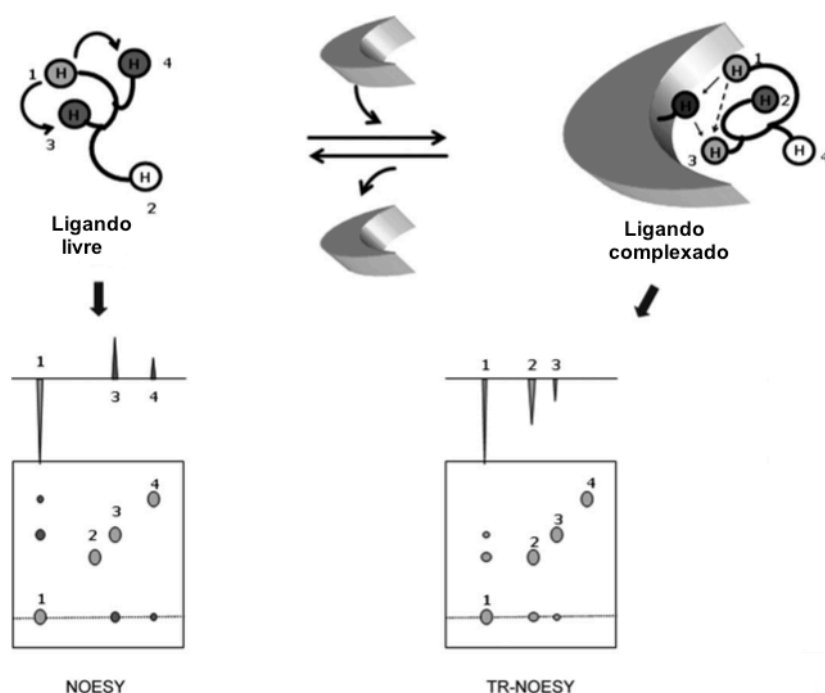


Figura I.10. Estudo da interação proteína-carboidrato, por trNOE. Na conformação livre do ligando o próton H1 está a interagir fortemente com o H3 e fracamente com o H4, após a ligação à proteína o ligando altera a sua geometria, passando a interagir fortemente com o próton H2 e fracamente com o próton H3. Após a ligação à proteína ocorre uma mudança de fase, com os sinais positivos a ficar com sinal negativo. Figura adaptada da referencia [85].

O último parâmetro que pode ser monitorizado está também ele associado à diferença de topologia e tamanho entre o ligando e a proteína é a medição de difusão em solução [87, 92-94]. Uma pequena molécula orgânica difunde cerca de 20 vezes mais rápido que um receptor molecular de 20kDa, assim na forma complexada a molécula de ligando irá apresentar uma taxa de difusão mais lenta. Todos os temas abordados até aqui no dizem respeito à técnica de RMN para elucidação de processos de reconhecimento molecular proteína-ligando encontram-se resumidas na tabela I.6.

Tabela I.6. Metodologias estabelecidas no estudo interação proteína-ligando. Experiências de RMN e parâmetros utilizados para estudar interações moleculares nos diferentes pontos de vista.

Abordagem	Molécula em estudo	Aplicações
Perturbação de desvios químicos	Proteína/ligando	Identificação dos resíduos em interação do ponto de vista da proteína ou ligando; constantes de afinidade.
STD-NMR	Ligando	<i>Screening</i> de ligandos, identificação dos epítomos de ligação
WaterLOGSY	Ligando	Identificação de compostos que interagem com a proteína.
Relaxação T_1 e T_2	Ligando/proteína	Identificação de regiões do ligando a interagir com o receptor molecular, através do relaxação estimulada pelo processo de reconhecimento. Permite estimar constantes de afinidade
TrNOE	Ligando	Dá informação acerca da geometria de ligação na sua forma ativa no local de ligação
Difusão	Ligando	Medição de diferentes taxas de difusão do ligando no estado livre vs. ligado

No âmbito, deste trabalho foram apenas estudados os processos de reconhecimento molecular, nos quais o ligando foi a molécula a ser observada. As metodologias utilizadas focaram-se em STD-RMN, trNOE e no estudo das perturbações dos desvios químicos ao longo de uma titulação.

I.5.1. STD-RMN- Transferência de saturação

O STD-RMN inicialmente desenvolvido por Meyer *et al* [88] faz uso do efeito nuclear de *overhauser* (NOE) para estudar processos de reconhecimento de pequenos ligando por parte de receptores biológicos. A amostra ideal para uma experiência de STD é constituída por uma proteína de médio a elevado peso molecular e um ligando de baixo peso molecular dissolvidos num tampão deuterado. O ligando é normalmente utilizado num grande excesso em relação à proteína à qual se liga num regime de permuta rápida, estando mais indicado para estudar a interação para ligandos de baixa afinidade.

O STD-RMN permite fazer estudos de *screening* de ligandos de um modo rápido e eficaz, demorando em média metade do tempo necessário para a aquisição de um espectro 2D-NOESY. Por outro lado, como o objectivo de técnica reside na “observação” do ligando, não são necessárias grandes quantidades de proteína (mM), nem mesmo que esta se encontre marcada isotopicamente. Nestes estudos, o receptor biológico apenas terá como função transferir a magnetização para o ligando. O STD-RMN permite por um lado determinar a ordem de grandeza da afinidade que o ligando tem para o receptor, e por outro lado fazer o mapeamento dos epítomos envolvidos no processo de reconhecimento molecular.

A experiência de ^1H STD-RMN envolve a aquisição de dois espectros. Um dos espectros 1D- ^1H denominado de *on-resonance* (I_{SAT}), no qual a proteína é irradiada com um campo magnético oscilante em ressonância com a frequência de Larmor. A irradiação é efectuada numa região do espectro onde se encontrem apenas sinais da proteína, nomeadamente em regiões blindadas entre os 0 e -1 ppm, nesta região encontram-se também os prótons alifáticos da proteína. O segundo espectro 1D- ^1H é adquirido, sem que o ligando ou a proteína sejam irradiados, sendo este denominado de *off-resonance*, cuja a intensidade dos sinais é dada por I_0 . Da subtração do espectro *off-resonance* e *on-resonance*, irá surgir um espectro no qual apenas irão estar presentes as ressonâncias do ligando que interagiu com a proteína, como é exemplificado na figura I.11.

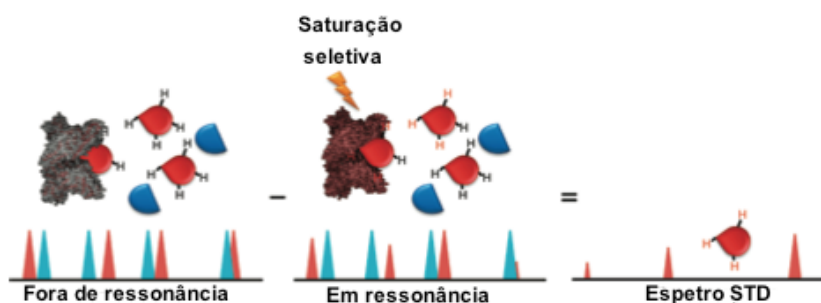


Figura I.11. Experiência de STD-RMN. A obtenção do espectro de STD (I_{STD}), resulta da diferença do espectro fora de ressonância (I_0 - referência) e do espectro no qual a proteína é seletivamente saturada (I_{SAT}). Figura adaptada da referência [95].

Assim, a técnica pode funcionar como um método de *screening* de ligandos, visto que num *pool* de ligandos, apenas irá aparecer no espectro final, as ressonâncias espectrais daqueles que

receberam magnetização da proteína (ou sejam estiveram em regime de NOE negativo com a proteína). Todos os restante ligandos, bem como contaminações irão desaparecer no espectro de diferença –impressão digital dos ligandos.

Capítulo II: Reconhecimento molecular de arabinoligómeros pela AraN

II.1. Introdução

A capacidade das bactérias prosperarem nos meios em que se inserem, está intimamente associada à aptidão que estas apresentam na captação de fontes energéticas. Para tal os microrganismos dispõem de um vasto e diverso conjunto de transportadores dedicados à aquisição nutrientes. Entre outros, a *Bacillus subtilis* utiliza uma SBP na captação inicial de oligossacáridos que contêm arabinose, associada a um transportador ativo primário do tipo ABC, constituindo o transportador AraNPQ-MsmX.

A análise *in vivo* destes sistema de transporte sugere fortemente que o transportador está associada à incorporação de oligómeros de arabinose (com duas a quatro subunidades) por parte da *Bacillus subtilis*, não reconhecendo a SBP a unidade monomérica de arabinose [9].

Com o objectivo de caracterizar o processo de reconhecimento molecular exercido por parte da AraN, recorreu-se à técnica de RMN para identificar quais as regiões do carboidrato que desempenham um papel ativo no processo de reconhecimento molecular e, adicionalmente caracterizar a afinidade deste processo recorrendo a STD-NMR. Por outro lado com recurso a ^1H - ^1H -NOESY (na ausência e presença de AraN), desvendar-se-á se o processo de reconhecimento molecular implica alterações conformacionais do açúcar após reconhecimento por parte da proteína.

Seguidamente utilizando uma abordagem bioinformática é gerado um modelo da AraN por homologia, bem como *dockings* moleculares e alinhamentos múltiplos (MSA) de elementos de estrutura secundaria com objectivo de validar resíduos putativos de se encontrarem a interagir com o ligando, do ponto de vista da proteína. Estes irão ser subsequentemente alvos de mutagénese. Finalmente é examinada a capacidade destes mutantes no reconhecimento molecular do ligando através da técnica de STD-RMN.

Em suma esta estratégia (figura II.1, página seguinte) permitirá estudar o processo de reconhecimento do ponto de vista do ligando e identificar resíduos da proteína envolvidos nesse processo.

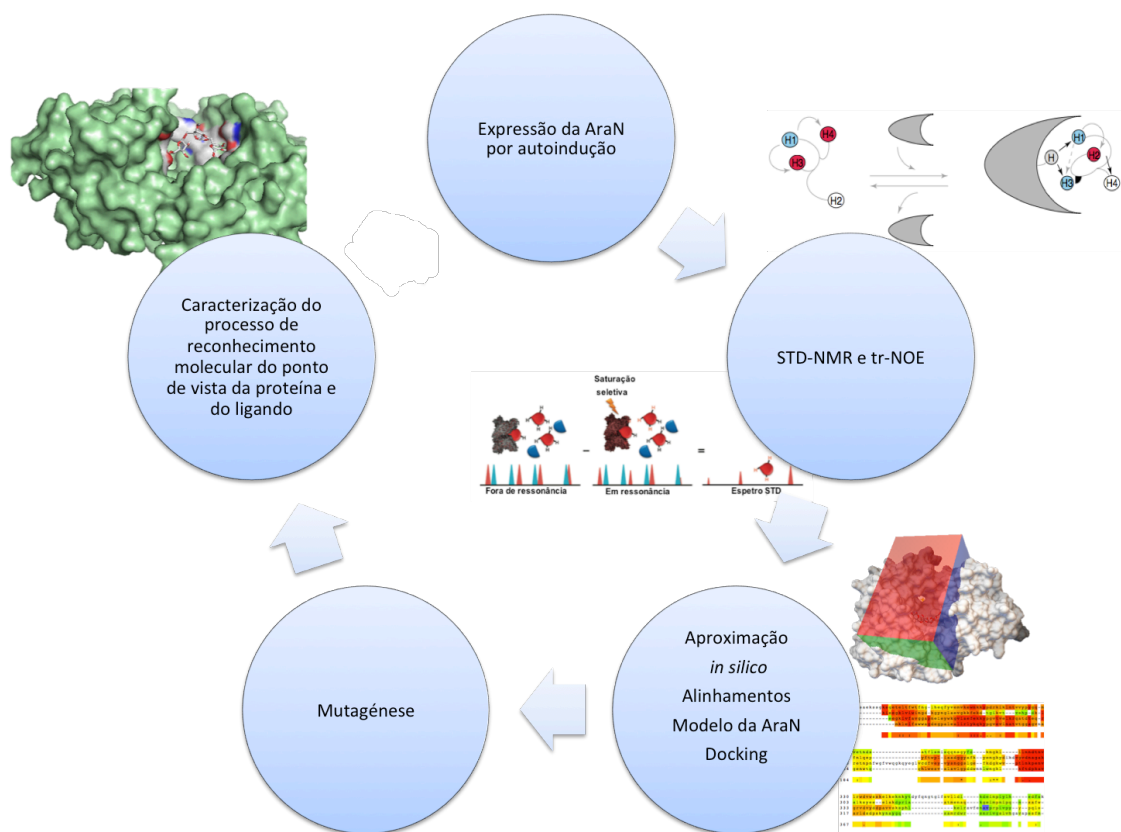


Figura II.1. Estratégia utilizada para estudar a interação protein-carbohidrato promovida por parte da AraN. Metodologias utilizadas para elucidação estrutural de um ponto de vista biomolecular dos processos de reconhecimento molecular entre a AraN e carbohidratos. Numa primeira fase a AraN é sobreexpressa por autoindução sem marcação isotópica, sendo empregues técnicas de RMN (STD-RMN e trNOE), para caracterizar o processo de reconhecimento molecular do ponto de vista do carbohidrato. Esta abordagem é complementada por estudos *in silico* (MSA e *docking* molecular) e de mutagénese dirigida para aferir quais os parceiros de reconhecimento por parte da AraN.

II.2. Materiais e métodos

II.2.1. Expressão heteróloga da AraN-His₆tag

A expressão da AraN-His₆tag foi efectuada a partir de um *stock* de glicerol de células competentes de *E.coli* BL21(DE3) pLysS (Novagen), transformadas com o vector pZI37, um derivado de pET30a (Novagen). Este vector, contém o gene *araN* sob o controlo do promotor T7, e permite a expressão de uma forma recombinante de AraN que possui uma cauda de histidina na região C-terminal (Inácio, J.M., e Sá-Nogueira, I., resultados não publicados).

Para a sobreexpressão da AraN, recorreu-se numa primeira fase ao plaqueamento das células competentes em ágar LB (1.6% agarose), contendo esta placa os antibióticos Cm (25 µg.mL⁻¹) e Km (20 µg.mL⁻¹).

De seguida, foi repicada uma colónia para 1L de meio de cultura de autoindução [96] (ver apêndice, secção VI.2. *Meios de cultura*), deixado a crescer durante 16h, à temperatura de 37°C e 180 rpm (Gallenkamp).

A proteína AraN que possui a substituição R81A foi produzida tal como acima descrito, utilizando células de *E.coli* BL21(DE3) pLysS (Novagen) transformadas com o plasmídeo pMJ36. Este vector é derivado de pZI37, e foi obtido por mutagénesse dirigida do alelo *araN* (Ferreira, M. J., e Sá-Nogueira, I., resultados não publicados).

II.2.2. Isolamento da AraN-His₆Tag

Numa primeira etapa recolhem-se as células por centrifugação (Sigma 3-16K Sartorius/ Rotor:19776) a 4°C durante 10 minutos a 7500 rpm. O sobrenadante é descartado e o extrato insolúvel (*pellet*) contendo as células é ressuspenso em tampão de fosfato de sódio 50 mM, 10 mM imidazole, 100 mM NaCl a pH=7,5, sendo também adicionado 1 µl de benzonase (Sigma-Aldrich), bem como 100 µl de PMSF (100 mM). De seguida foi efectuada a lise celular com o recurso a um sonicador (Hielscher UP50H/UP100H ponta MS7) em ciclos de amplitude 70%, durante 30 segundos, com 30 segundos de intervalo entre cada ciclo, sendo este processo repetido durante 7 ciclos.

Obtido o lisado, este foi submetido a uma segunda centrifugação com intuito de isolar as proteínas sobreexpressas dos restantes detritos celulares. Neste passo procedeu-se à centrifugação durante 30 min a 9500 rpm. Finalmente o sobrenadante é filtrado com recurso a um filtro de 0,45 µm, e mantido a 4°C para subsequente purificação.

II.2.3. Purificação da AraN-His₆tag

Posteriormente ao isolamento da proteína, procedeu-se à purificação da AraN-His₆tag. Para tal recorreu-se a uma cromatografia em coluna de afinidade por imobilização através de ião metálico (GE Healthcare Life Sciences). Inicialmente a coluna (1mL de volume) é lavada com 5 volumes de água, sendo subsequentemente carregada com 2 volumes de 100mM NiSO₄. De seguida a coluna é novamente lavada com 5 volumes de água para remover o níquel que não fica complexado à matriz de agarose funcionalizada com NTA, sendo finalmente equilibrada com 10 volumes de tampão. Preparada a coluna, esta é carregada com o sobrenadante previamente filtrado (secção II.2.2.), a um

caudal de 1 ml/min. Após carregado o sobrenadante, a coluna é novamente lavada com 10 volumes de tampão antes de se proceder à eluição das proteínas. A eluição foi efectuada através da passagem de 5 volumes do tampão de eluição (50mM de Fosfato de sódio e 100mM de NaCl) com diferentes forças iónicas de imidazole (40mM a 500mM). A passagem de tampão com 40mM de imidazole serve para eluir proteínas que apresentem resíduos de histidina à sua superfície, para que de seguida a eluição da AraN-His₆tag, nas fracções de 100mM e 200mM de imidazole, seja efectuada com um elevado grau de pureza.

Para confirmar em que fracções foi eluída a proteína e o qual o seu grau de pureza efectuou-se uma SDS-PAGE (sendo a percentagem de acrilamida de 12,5%). Para tal é necessário proceder-se à preparação do gel de corrida, bem como do gel de deposição.

Para análise por SDS-PAGE prepararam-se amostras de fracções recolhidas durante todo o processo de isolamento e purificação da proteína (sobrenadante, *pellet*, *flow through*, lavagens e fracções de purificação). As amostras são preparadas pela adição de 8 µl de cada Fração e 2 µl de *loading buffer*. De seguida estas amostras são incubadas durante 5 minutos num banho seco (Fischer Scientific FB 15101) a 100° C, e finalmente são rapidamente centrifugadas (Sorvall Heraeus) (≈5 segundos) para ficarem homogéneas e recuperarem parte da água que condensa no banho seco. O SDS-PAGE, foi efectuada com uma intensidade de corrente eléctrica constante de 30 mA por gel. O gel resultante encontra-se na secção II.3.1. *Expressão heteróloga e purificação da AraN* nas figuras II.3 e II.4.

As fracções de eluição puras são então combinadas numa só. Com o objectivo de remover o imidazole, a amostra correspondente às fracções combinadas é novamente purificada numa coluna PD-10 Sephadex G-25 Medium (GE Healthcare). Inicialmente a coluna é carregada com 25mL de tampão, sendo de seguida carregada com 2,5mL de amostra, com o *outflow* a ser desprezado. Finalmente a coluna é carregada com 3,5mL de tampão (50mM de PBS sódio, 100mM NaCl) sendo a proteína eluída neste passo. Este processo foi repetido duas vezes. Obteve-se um volume de 7mL que foi quantificado pelo método de Bradford (II.2.4.) De seguida, a amostra foi dividida em alíquotas de volumes de 1mL e 450µL tendo sido posteriormente liofilizadas e armazenadas a -20°C.

II.2.4. Quantificação da AraN-His₆tag

Para determinar a concentração da AraN-His₆Tag obtida em II.2.3 utilizou-se o método de Bradford (Bio-Rad Protein Assay). Neste método, é utilizado o corante Coomassie Brilliant Blue G-250 que pode apresentar uma cor vermelha que absorve a 465nm, ou pode ainda apresentar uma cor azul absorvendo a 595nm. A cor vermelha do corante é transformada em azul quando ocorre ligação do corante à proteína em estudo, sendo proporcional o aumento de absorvância à quantidade de proteína presente [97]. Para o passo de quantificação a AraN *wild-type* foi diluída num factor 1:50 apresentando uma absorvância de 0,282, ao passo que a R81A foi sujeita a uma diluição com um factor de 1:10 (consultar secção V.3. *Quantificação pelo método de Bradford*). Ambas as medições foram efectuada em pontos triplicados.

Relativamente à AraN *wild-type* foi obtido um rendimento de 8,8 mg/L cultura, ao passo que o rendimento do mutante R81A foi de 9 mg/L cultura (em ambos os casos apenas foi quantificada a

proteína pura, portanto este rendimento não pode ser considerado um rendimento global de sobreexpressão).

II.2.5. Preparação dos carboidratos

Todos os carboidratos (arabinobiose, arabinotriose, arabinotetraose, arabinopentaose, celobiose, celotetraose, xilopentaose, maltohexaose e manotriose) foram adquiridos comercialmente à empresa Megazyme (Country Wicklow, Ireland) apresentando uma pureza de $\approx 95\%$ tendo sido utilizados em solução de DMSO, sem nenhuma purificação adicional. A arabinose foi por seu turno adquirida à empresa Sigma-Aldrich com um grau de pureza de $\approx 99\%$, também ela preparada em solução de DMSO.

II.2.6. Aquisição de espectros de RMN

Todos os espectros para a caracterização dos carboidratos em estudo e para estudos de interação por STD-RMN foram adquiridos num espectrofotómetro de 600 MHz Bruker Avance III, equipado com uma sonda criogénica (CP-TCI) operando a uma frequência de 600.13 MHz em ^1H , bem como num espectrofotómetro de 400 MHz Bruker Avance III, equipado com uma sonda TXI, operando a uma frequência de 400.15 MHz em ^1H . Foram utilizadas sequências de pulso presentes na biblioteca de sequências do software Bruker Topspin 3.1. A temperatura de aquisição dos espectros foi de 298 K excepto quando indicada outra temperatura. Todos os dados foram processados com o programa Bruker TopSpin 3.2.

II.2.6.1. Caracterização de carboidratos

Alguns ligandos em estudo foram caracterizados (secção II.3.2) por ^1H e ^{13}C RMN (arabinobiose, arabinotriose, arabinotetraose). A atribuição das ressonâncias espectrais dos carboidratos foi conseguida, com auxílio de espectros ^1H , COSY, TOCSY, 1D-selectTOCSY, HSQC e HSQC-TOCSY, bem como por comparação com dados da literatura *Perez et al* [98]. A caracterização da celobiose a celotetraose foi retirada da dissertação de doutoramento de Aldino Viegas [99].

O espectro de ^1H (sequência zg) foi adquirido numa janela espectral de 12335,53 Hz (20,555 ppm) centrada a 3706,05 Hz (6,175 ppm) tendo sido adquiridos 16 scans, com 64K pontos e um tempo de relaxação de 1,0 segundo. O mesmo espectro com a supressão água utilizando a técnica de supressão de água por *excitation sculpting* (sequência zgesgp) foi efectuado, com auxílio de um pulso seletivo Sinc1.1000 a 2824,20 Hz (4,706 ppm) durante 2 ms.

O espectro COSY (sequência cosygpqf) foi adquirido com 8 scans por incremento numa matriz de 2048 pontos em F2, com uma janela espectral de 2000,00 Hz (5,000 ppm), centrada a 1880,56 Hz (4,700 ppm) e 256 incrementos em F1, com um tempo de relaxação de 1,5 segundos.

O espectro HSQC (sequência hsqcetgpsisp2) foi adquirido com 16 scans por incremento com 1024 pontos em F2, numa janela espectral de 4795,396 Hz (7,991 ppm), centrada a 2824,09 Hz (4,706 ppm), tendo sido feitos 256 incrementos em F1 numa janela espectral de 12073,13 Hz (80,000 ppm), com um tempo de relaxação de 1,5 s. Neste espectro foi utilizado um tempo de evolução de 1,724 ms ($J_{\text{CH}}=145$ Hz).

Relativamente ao espectro HSQC-TOCSY (sequência hsqcdiedetgpsisp.1), este foi adquirido com 16 scans numa matriz de 1024 pontos na dimensão F2, com uma janela espectral de 4795,396 Hz (7,991 ppm), centrada a 2824,09 Hz (4,706 ppm), com 256 incrementos em F1 numa janela espectral de 12073,069 Hz (80,000 ppm), com um tempo de relaxação de 1,5 s. Tal como na aquisição do espectro HSQC, neste foi calculado um tempo de evolução para correlações ^1H - ^{13}C de 1,724 ms ($J_{\text{CH}}=145$ Hz). Na aquisição destes espectro foi utilizado um tempo de mistura de 45 ms.

Finalmente, o selTOCSY-1D (sequência selmlgp), foi adquirido com 16 scans numa janela espectral de 12335,53 Hz (20,555 ppm) centrada no sinal de interesse e irradiado seletivamente, com 64K pontos e um tempo de relaxação de 1s. A utilidade da técnica reside na capacidade de irradiar seletivamente sinais pertencentes a diferentes subunidades de açúcar, deixando evoluir a magnetização para os prótons acoplados escalarmente, durante tempos de mistura diferentes, permitindo assim efetuar uma atribuição sequencial dos prótons de um sistema de spin entre vários no mesmo oligossacárido. A irradiação seletiva foi conseguida através de um pulso seletivo de 180°, Gauss1_180r.1000 durante 80 ms, centrado nas diferentes frequências dos prótons anóméricos constituintes do açúcar. Os tempos de mistura foram de 40, 60, 80 e 100 ms.

II.2.6.2. Estudos de STD-NMR

Para a aquisição dos espectros de STD-RMN, foram utilizados os programas de pulso com ou sem supressão de água, presentes na biblioteca Bruker: i) stddiff a 298K para otimizar o tempo de saturação (t_{sat}) (secção II.3.3.2, durante o qual a proteína é irradiada, bem como na titulação da AraN com arabinotriose para determinar a K_d (secção II.3.3.3) ii) stdiffesgp.3 para fazer o mapeamento dos epítomos de ligação (secção II.3.3.4 e II.3.3.5).

As experiências relativamente a optimização de t_{sat} foram executadas com recurso ao programa de pulso stddiff (sem supressão do sinal residual da água), foram adquiridas no espectrofotómetro de 600 MHz com 16 scans x 32 repetições, tendo sido adquiridos 14336 pontos numa janela espectral de 9615,4 Hz (16,0221 ppm) centrada a 2823,65 Hz (4,705 ppm). Foram adquiridos espectros com diferentes tempos de saturação 0.1, 0.2, 0.4, 0.6, 0.8, 1, 1.20, 1.40, 1.60, 1.80, 2, 2.50, 3 e 3.50 s. Na experiência *on-resonance* a proteína foi irradiada a -150 Hz, com recurso a um pulso seletivo Eburp2.1000 (pulso de 90° de 50ms de duração). A experiência *off-resonance* foi adquirida com a irradiação a 20000 Hz. Para estes estudos analisou-se uma mistura composta por 8,9mM de arabinotriose e 58,0 μM de AraN, num excesso de 1:154 de ligando. A mistura encontra-se dissolvida em 92,6% de D_2O e 7,4% de DMSO.

Os espectros de 1D STD-RMN foram obtidos pela subtração do espectro *on-resonance* ao espectro *off-resonance*.

Ao longo da titulação do proteína com ligando, para o calculo do K_d a concentração de proteína foi mantida a 25 μM , sendo esta titulada com ligando até um excesso de ligando:proteína de 500:1. Inicialmente são preparadas duas soluções stock, com iguais concentrações de AraN (25 μM) no entanto uma apresenta um excesso de 1:10 de ligando (241,3 μM de arabinotriose) e a outro um excesso de 1:500 de ligando (12,5 mM de arabinotriose). A mistura encontra-se dissolvida em 89,6% de D_2O e 10,4% de DMSO.

Na tabela II.1 estão indicados os pontos da titulação, sendo que o volume adicionada é proveniente da solução 1:500 (que titula a solução 1:10). As condições de aquisição dos espectros de STD-RMN foram idênticas às anteriormente descritas para a otimização do tempo de saturação com um tempo de saturação fixo em 1,5 segundos

Tabela II.1. Titulação AraN-arabinotriose. Tabela com os pontos utilizados na curva de titulação em STD-RMN, para a determinação do K_d da arabinotriose na presença da AraN.

[P] μ M	[L] μ M	[L]/[P] eq	Volume
25,1	241,3	9,6	0
25,1	482,7	19,3	10
25,1	938,0	37,4	20
25,1	1360,3	54,3	20
25,1	1752,9	69,9	20
25,1	2623,6	104,7	50
25,1	3364,3	134,2	50
25,1	4557,2	181,8	100
25,1	5860,3	233,7	150
25,1	6395,4	255,1	80

Nos espectros recolhidos para o mapeamento de epítomos de ligação a solução foi preparada com uma razão proteína:ligando de 1:100 (2,5mM de carbohidrato e 25 μ M de AraN), estando esta solução em 93,7% de D₂O com 6,3% de DMSO. Todas as experiências, salvo alguns casos devidamente identificados, foram efectuadas a 283K (na secção II.3.3.2. Efeito da temperatura foram adquiridos espectros de STD com diferentes temperaturas). As experiências de STD-RMN foram adquiridas utilizado o programa de pulso stddiffesg para os seguintes carbohidratos: arabinose, arabinobiose, arabinotriose (estas experiências foram adquiridas a diferentes temperaturas: 283K, 298K e 308K), arabinotetraose, arabinopentaose, celobiose, celotetraose, maltohexaose, manotriose e xilopentaose. Para todas estas experiências as condições foram idênticas às anteriormente utilizadas excepto que se adquiriram 8 scans com 32 incrementos e um tempo de saturação de 1,5 segundos.

Relativamente aos estudos de competição (secção II.3.3.7) estes foram efectuados entre i) arabinotriose e maltohexaose, ii) arabinotriose e celotetraose. No primeiro caso foi adquirido um espectro de STD-NMR de uma amostra composta por AraN a 25 μ M, arabinotriose 1,25 mM e maltohexaose 1,25 mM (os carbohidratos encontravam-se na proporção estequiométrica 1:1, num excesso total de 1:100 relativamente à proteína). Por seu turno, no segundo caso foram adquiridos três espectros nas quais os carbohidratos estavam em diferentes proporções estequiométricas, com a AraN a uma concentração constante de 25 μ M, e um excesso total de ligando de 1:100 relativamente à proteína nas seguintes proporções : i) proporção estequiométrica de 1:1 com 1,3 mM de arabinose e celotetraose ii) excesso 1:2 de arabinotriose (1,7 mM) sobre a celotetraose (0,8 mM), e finalmente iii) excesso de 1:2 de celotetraose (1,7 mM) sobre a arabinotriose (0,8 mM).

II.2.7. NOE e trNOE

O espectro de NOESY (secção II.3.8) (noesyegpph.rak) foi adquirido com 16 scans por incremento numa matriz de 2048 pontos em F2, com uma janela espectral de 3000 Hz (5,0 ppm), centrada a 2821,9 Hz e 512 incrementos em F1, com uma janela espectral de 3000 Hz (5,0 ppm), centrada a 2821,9 Hz. O espectro NOE sem proteína foi adquirido com o tempo de mistura de 400 ms, ao passo que o trNOE foi obtido com um tempo de mistura de 200 ms. Neste último a proteína estava num excesso de 1:10 relativamente ao ligando.

II.2.8. Modelo de homologia estrutural 3D da AraN

A construção do modelo de homologia estrutural 3D da proteína AraN (secção II.3.4) foi efectuado no servidor I-TASSER (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>) [100]. Foi submetida a seguinte sequência:

MCSAEKSSGKSGETELTFWTFNGLHEQFYVEMVKEWNKKYPDRKIKLNTVVYPYGQMHDNLSISLIA
GEGVPDIADVELARFSNFLKGSDIPLADLTPLIEKDRDKFVEARLTLYSKNGKLYGLDTHVGTTFMFYN
MDVMKKAGVNPDDIKTWDDYHKAGQKVRKVTGKPMGTVETNDSATFLSMISQQNSGYFDKNGKLIL
NNDTNVKTLYLKDMINDKTMIPAPGGGHHSEEYYGFMNQGAASVLMPIWYMGRFIDYMPDLKGKI
AIRPLPAWKEGGDRSAGLGGTATVVPKQSKHVELAKEFLAFAKGSEEGNKKLWSVLGFDPLRWDVW
SSKELKEKNKYTDYFQNGTGIFSVLLDIKDEINPIYLHEDFAKASDLVNRSVLFDAKSSQQKTPKQALDR
AAGELKQKHLE

O melhor modelo obtido (secção, II.3.4.1. *Modelo de homologia estrutural 3D da AraN*) foi construído seguindo um protocolo bem estabelecido de *Zhang et al* [101]. Este modelo foi utilizado para prosseguir com estudos de modelação molecular de *docking* com os carboidratos.

II.2.9. Docking molecular

Para efetuar os *dockings* moleculares (secção II.3.6) foi utilizado o programa Autodock 4.2 [102]. Como receptor molecular foi utilizado o modelo calculado no I-TASSER. Os ligandos utilizados foram a arabinobiose [PDB:2X8S] [103] e arabinotriose.[PDB:2W5O] [104]. Estes ligandos encontram-se ambos na conformação α .

A preparação das estruturas do receptor e do ligando para os estudos de *docking* envolveu previamente a adição de átomos de hidrogénios polares ao receptor e o cálculo de cargas (cargas kollman para o receptor e gasteiger para o ligando).

Nos *dockings* efectuados, foram utilizadas 10 simulações LGA (*Lamarckian Genetic Algorithm*). O tamanho da amostra foi de 150, o numero máximo de avaliações foi definido para 2,5E6 e o numero máximo de gerações para 2,7E4. Entre cada processo iterativo foi definido que só sobrevivia um individuo (conformação de mais baixa energia). A taxa de mutações genéticas foi de 0,02 e a taxa de cruzamento 0,8. Foi definida uma caixa centrada em X=-12.778, Y=1.861 e Z=0.22, de dimensões 66X126X80 (X,Y,Z) com um espaçamento de 0.375 Å.

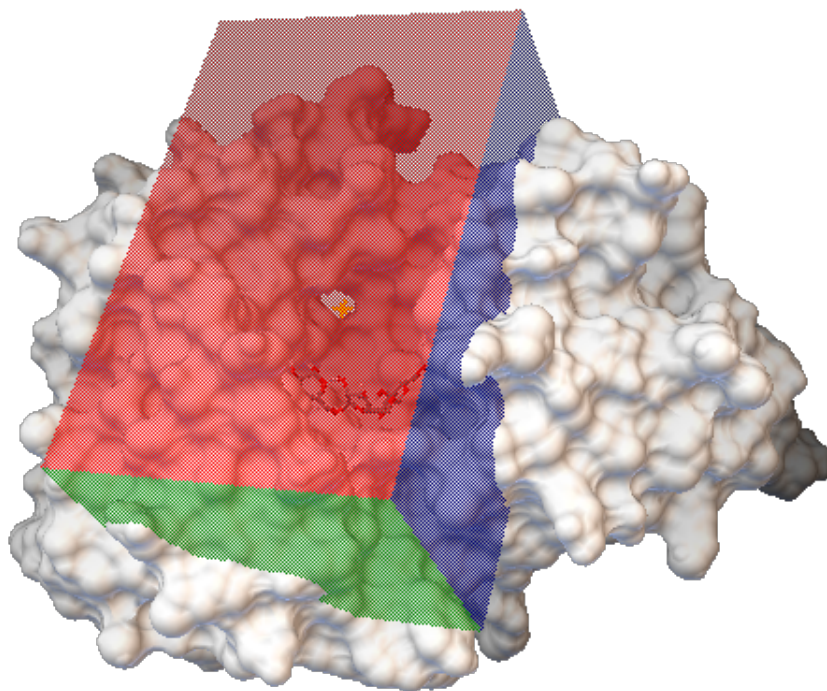


Figura II.2. Gridbox utilizada no docking molecular. Imagem ilustrativa do posicionamento da *gridbox* ao longo do local de ligação da AraN.

As 10 conformações finais foram organizadas num *cluster* de conformações com um *cutoff* de 2,0 Å.

II.2.10. Análise de homologia da sequência da AraN com outras SBP

A construção de alinhamentos da AraN com outras SBP's (secção II.3.5) cujas estruturas tridimensionais se encontram depositadas no PDB foi efectuada com a ferramenta T-Coffee cujo o protocolo é descrito por *Notredame et al* [105, 106]. Relativamente à ferramenta T-Coffee, a modalidade adoptada para computar os alinhamentos foi a Expresso [105]. Este modelo é o que permite i) obter um maior grau de precisão ii) efetuar um alinhamento múltiplo de sequências tendo em conta propriedade estruturais dos padrões utilizados [106].

Assim e com o intuito de identificar resíduos putativos de reconhecer a arabinotriose no caso da AraN foram efectuados alinhamentos (ver secção Anexos, Alinhamentos) com as seguintes sequências de proteínas: Maltose BP de *E.coli* (PDB: 3MBP) [107], ciclo/maltodextrina BP de *Thermoactinomyces vulgaris* (PDB:2ZYO) [108], galacto-N-biose/lacto-N-biose BP de *Bifidobacterium longum* (PDB:2Z8D) [109], maltose BP de *Thermus thermophilus* (PDB:2GH9) [110], GacH BP de *Streptomyces glaycescens* (PDB:3K00) [111], trehalose/maltose BP de *Thermococcus litoralis* (PDB:1EU8) [112], maltodextrin BP de *Pyrococcus furiosus* (PDB:1ELJ) [113], maltooligosacarido BP de *Streptococcus pneumoniae* (PDB:2XD3) [54] e da glucose BP de *Thermus thermophilus* (PDB:2B3B) [114].

II.3. Resultados e discussão

II.3.1. Expressão heteróloga e purificação da AraN

A sobreexpressão da proteína AraN ($M_w=47\text{kDa}$) pode ser confirmada pela presença de uma banda com 47kDa no gel de SDS-PAGE como é possível constatar na figura II.4 nos poços de 3 a 8.

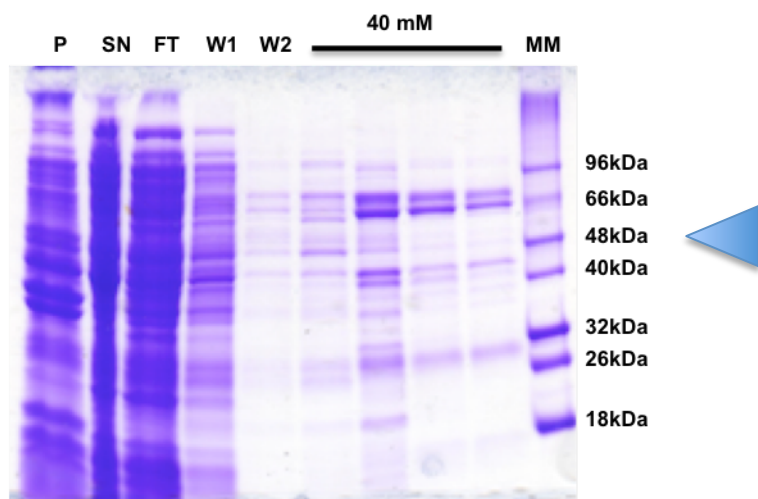


Figura II.3. SDS-PAGE (12,5%), com as fracções de purificação da AraN. Poço 1- Fração insolúvel (*pellet*). Poço 2- Fração solúvel (sobrenadante). Poço 3- *Flow through*. Poço 4- *Wash* 1. Poço 5- *Wash* 2. Poço 6 a 9- Fração eluição 40mM imidazole. Poço 10- Marcador de pesos moleculares. $M_w(\text{AraN})=47\text{kDa}$.

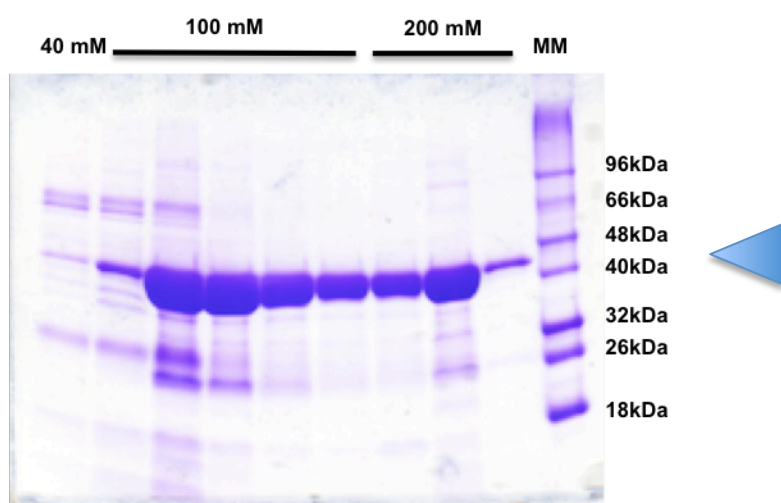


Figura II.4. SDS-PAGE (12,5%), com as fracções de purificação da AraN. Poço 1- Fração eluição 40mM. Poço 2 a 6- Fração eluição 100mM imidazole. Poço 7 a 9- Fração eluição 200mM imidazole. Poço 10- Marcador de pesos moleculares. $M_w(\text{AraN})=47\text{kDa}$.

Com o objectivo de estudar a interação da AraN com diferentes carboidratos, foram escolhidas apenas as fracções mais puras, para prevenir a eventual interação com outras proteínas que não a AraN. Para tal escolheram-se a quarta e quinta fracções de 100 mM imidazole, bem como a primeira fração de 200 mM de imidazole (presente no segundo gel SDS-PAGE, nos poços 5, 6 e 7 respectivamente). O imidazole e outros contaminantes de baixo peso molecular foram retirados com recurso a uma cromatografia em coluna PD10 (secção II.2.3.). No passo de quantificação a AraN

wild-type foi diluída num factor 1:50 apresentando uma absorvância de 0,282 (secção VI.3). Ambas as medições foram efectuadas em pontos triplicados.

Relativamente à AraN *wild-type* foi obtido um rendimento de 8,8 mg/L de cultura (neste caso apenas foi quantificada a proteína pura, portanto este rendimento não pode ser considerado um rendimento global de sobreexpressão).

II.3.2. Caracterização de carboidratos

A atribuição inequívoca de todas, ou do máximo de ressonâncias dos espectros 1D- ^1H RMN dos carboidratos em estudo é um procedimento obrigatório para possibilitar a análise dos estudos de ^1H STD-RMN com vista ao posterior mapeamentos dos epítomos que se encontram a interagir com a AraN. Essa atribuição é conseguida através de um conjunto de experiências de ^1H , ^{13}C - ^1H -HSQC, ^{13}C - ^1H -HSQC-TOCSY, 1D-seITOCY (secção II.2.6.1), bem como pela recurso a dados disponíveis na literatura para alguns dos carboidratos em estudo [98, 99].

II.3.2.1. Atribuição de ressonâncias espectrais de carboidratos

O resultado da atribuição das ressonâncias espectrais da arabinobiose, arabinotriose, arabinotetraose, celobiose e celotetraose encontra-se resumido nas figuras II.7 a II.11, respectivamente.

Tendo em conta que o procedimento utilizado na atribuição das ressonâncias dos carboidratos da família dos arabinanos (arabinobiose, arabinotriose, arabinotetraose) foi em tudo semelhante, apenas irei apenas exemplifica-lo em seguida para a atribuição da arabinobiose. Relativamente aos celoligossacáridos, estes tinham sido previamente caracterizados [99].

Em solução a *L*-arabinobiose encontra-se na conformação de furano, distribuída numa mistura de tautómeros α/β numa razão de 60/40 respectivamente, como é ilustrado na figura II.5. Os sinais dos prótons anoméricos apresentam desvios químicos muito característicos, nomeadamente os das formas α e β da unidade redutora. Uma vez identificado um sinal de um próton anomérico para uma determinada unidade do oligómero os restantes sinais pertencentes ao mesmo sistema de spin (conjunto de prótons acoplados em cascata) podem ser facilmente identificados através dos espectros ^1H , ^{13}C -HSQC-TOCSY e 1D-seITOCY. Estes sinais constituem assim na maioria das vezes o ponto de partida para a atribuição de todos os outros sinais do espectro.

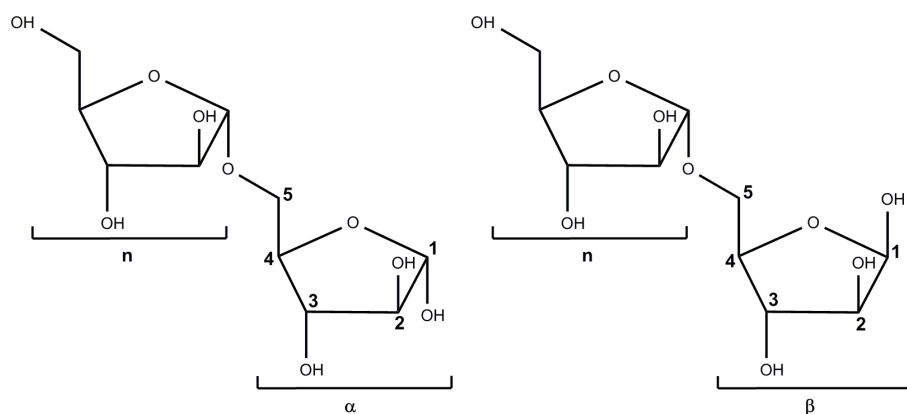


Figura II.5. Epimeros de arabinobiose. Estrutura de ambos os anómeros de arabinobiose, com a numeração dos respectivos prótons.

Através da sobreposição do espectro de $^1\text{H}, ^{13}\text{C}$ -HSQC com o espectro de $^1\text{H}, ^{13}\text{C}$ -HSQC-TOCSY tal como exemplificado na figura II.6, foi possível identificar os sistemas de spin, referentes aos dois monómeros. Como na sequência de TOCSY a magnetização só é transferida dentro do mesmo sistema de spin não passando através da ligação glicosídica para sistemas de spin adjacentes o espectro de HSQC-TOCSY permite identificar facilmente as correlações por acoplamento escalar entre os núcleos do mesmo sistema de spin. De referir ainda que nesta técnica foi utilizada uma modulação de fase por *J-coupling*, que permitiu também ela fazer uma diferenciação entre os grupos CH's (azul) e CH₂'s (verde), como é evidenciado na figura II.6.

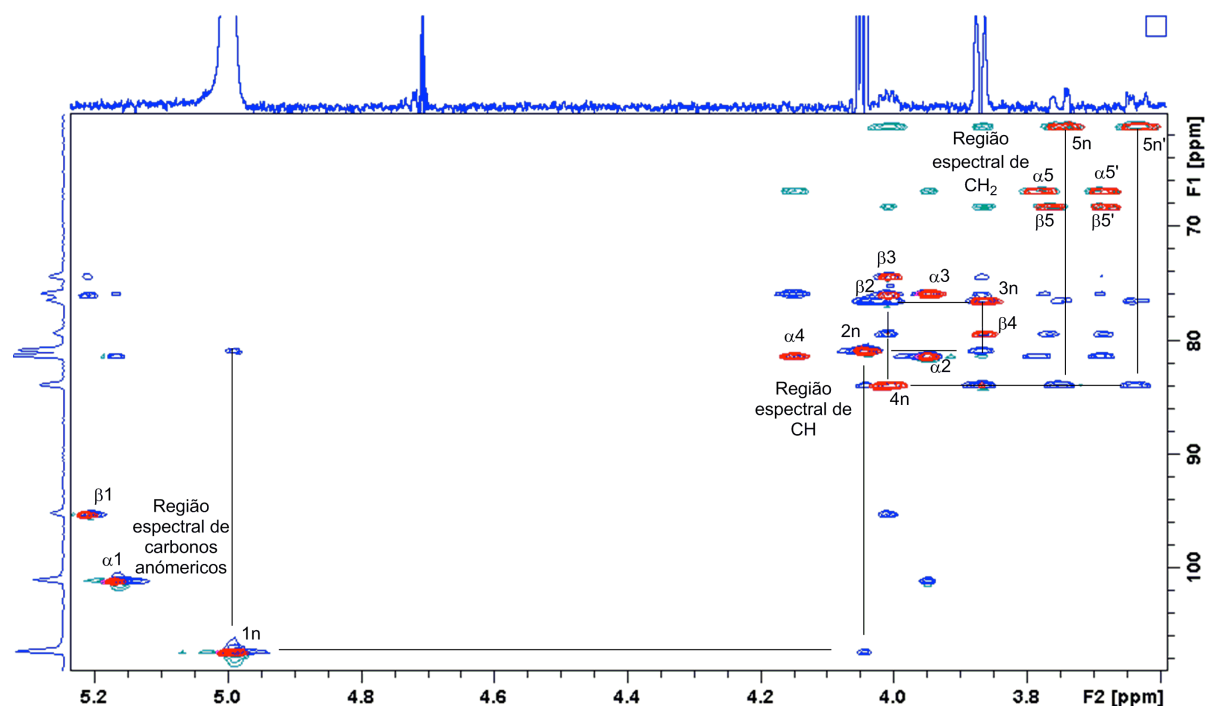


Figura II.6. Atribuição espectral da arabinobiose. A vermelho encontra-se representado o espectro ^{13}C - ^1H -HSQC. O espectro ^{13}C - ^1H -HSQC-TOCSY, apresenta dois contornos: contornos a azul representam CH's ao passo que contornos a verde representam CH₂'s. Na mesma figura encontra-se representada na dimensão F2 uma projeção externa de um espectro 1D-selTOCSY com O₁ centrado no protão anomérico da subunidade não redutora da arabinobiose. Através das correlações TOCSY foi também exemplificada a atribuição sequencial do sistema de spin referido anteriormente.

Para fazer a atribuição sequencial de ressonâncias para o mesmo sistema de spin foi também utilizada a técnica de 1D-selTOCSY. Para tal é necessário irradiar seletivamente a frequência do protão anomérico do sistema de spin que se pretende atribuir. Através da variação do tempo de mistura, esta magnetização irá sendo transferida do protão anomérico para os protões adjacentes. Para tempos de mistura curtos, é possível estabelecer correlações a curtas distâncias $^3J_{\text{CH}}$ (semelhante ao COSY), e para tempos de mistura mais longos a magnetização irá ser transferida para todos os protões dentro de uma mesmo sistema de spin. Visto que o anel de cada monómero de arabinose fecha com um oxigénio (proveniente da ciclização da aldose para gerar o hemicetal), para tempos de mistura curtos é possível estabelecer correlações entre o protão anomérico e o protão H2, sem que esta magnetização passe para o protão H4 ou H5. Na figura II.6 a na dimensão F2 encontra-se projetado o espectro 1D-selTOCSY irradiado a 4.995 ppm's sobre o

protão anomérico H1n com tempo de mistura de 120ms, permitindo a evolução da magnetização sobre todo o mesmo sistema de spin com a atribuição de ressonâncias dentro deste sistema de spin.

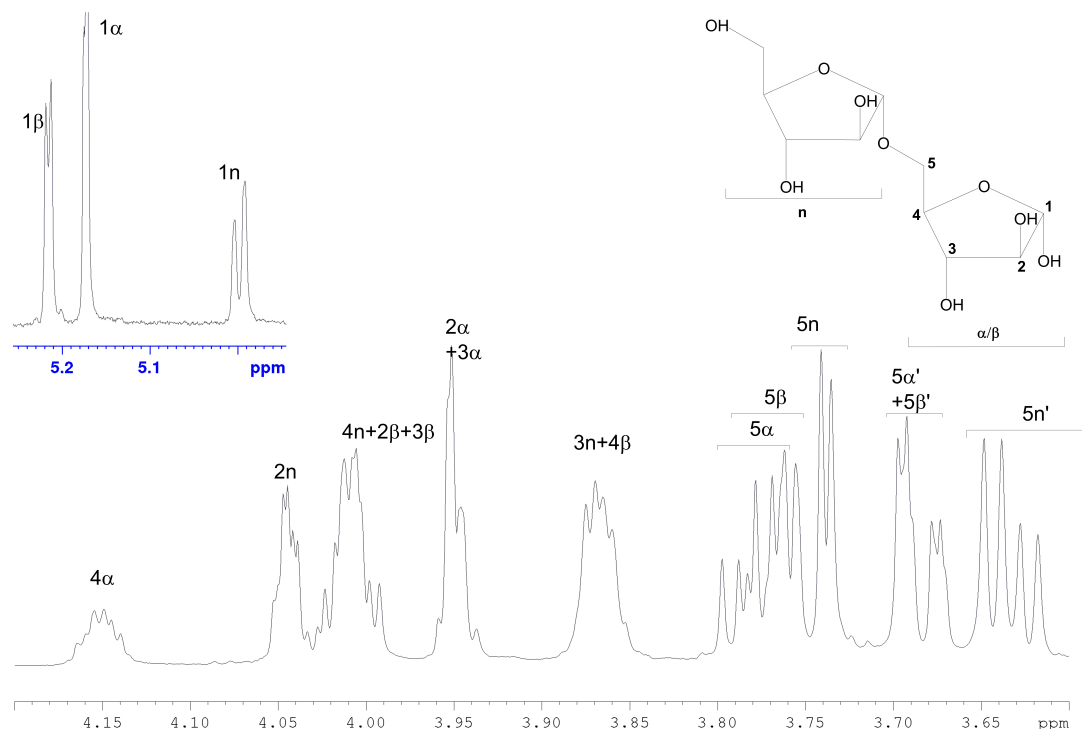


Figura II.7. Estrutura e atribuição do espectro de ^1H -RMN da arabinobiose. O espectro foi adquirido com uma solução de 2,5 mM de arabinobiose num espectrofotômetro de 600 MHz a 298 K com 32 scans.

Tabela II.2. Desvios químicos ^1H da arabinobiose. Tabela com as atribuições das ressonâncias espectrais (ppm) ^1H da arabinobiose.

Sistema de spin	H1	H2	H3	H4	H5	H5'
α	5,172	3,951	3,952	4,149	3,774	3,681
β	5,213	4,009	4,009	3,863	3,764	3,675
n	4,992	4,044	3,866	4,007	3,738	3,636

Tabela II.3. Desvios químicos ^{13}C da arabinobiose. Tabela com as atribuições das ressonâncias espectrais (ppm) ^{13}C da arabinobiose.

Sistema de spin	C1	C2	C3	C4	C5	C5'
α	101,095	81,310	75,819	81,287	66,800	66,800
β	95,248	75,954	74,333	79,353	68,150	68,150
n	107,342	75,954	76,443	83,879	61,150	61,150

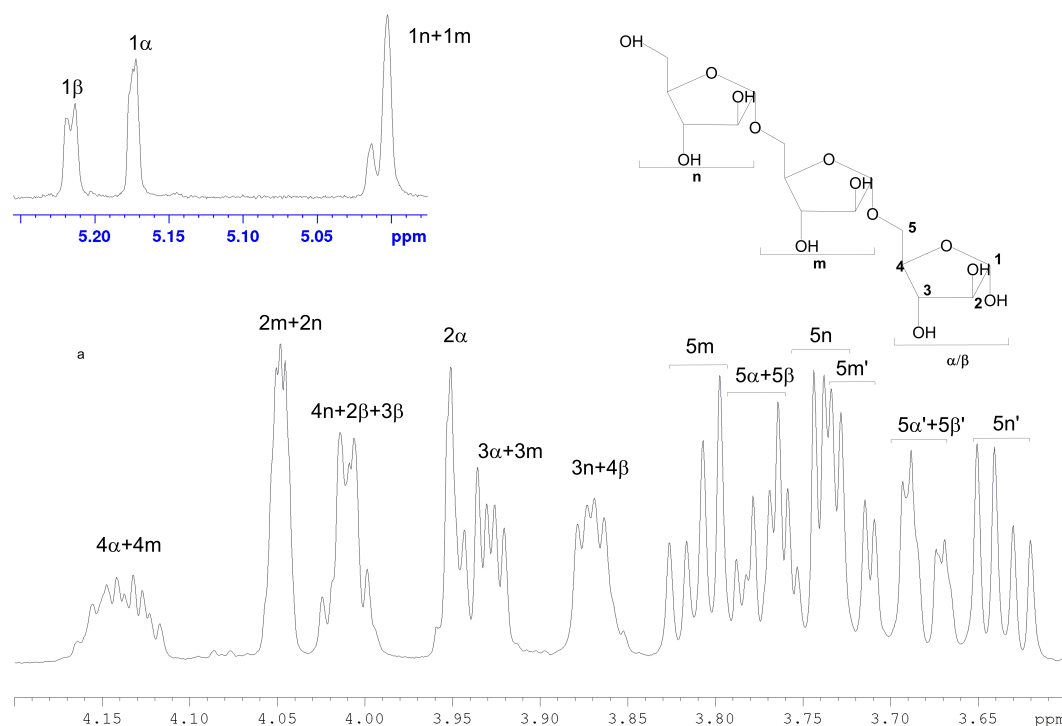


Figura II.8. Estrutura e atribuição do espectro de ^1H -RMN da arabinotriose. O espectro foi adquirido com uma solução de 2,5 mM de arabinotriose num espectrofotômetro de 600 MHz a 298 K com 32 scans.

Tabela II.4. Desvios químicos ^1H da arabinotriose. Tabela com as atribuições das ressonâncias espectrais (ppm) ^1H da arabinotriose.

Sistema de spin	H1	H2	H3	H4	H5	H5'
α	5,161	3,945	3,945	4,135	3,768	3,673
β	5,203	4,010	4,004	3,857	3,758	3,673
m	4,993	4,038	3,912	4,122	3,790	3,712
n	4,993	4,038	3,865	3,999	3,742	3,634

Tabela II.5. Desvios químicos ^{13}C da arabinotriose. Tabela com as atribuições das ressonâncias espectrais (ppm) ^{13}C da arabinotriose.

Sistema de spin	C1	C2	C3	C4	C5	C5'
α	101,040	81,327	75,830	81,270	66,890	66,890
β	95,225	75,940	74,330	79,332	68,247	68,247
m	107,320	80,790	76,660	82,270	66,747	66,747
n	107,32	80,790	76,434	83,830	61,100	61,100

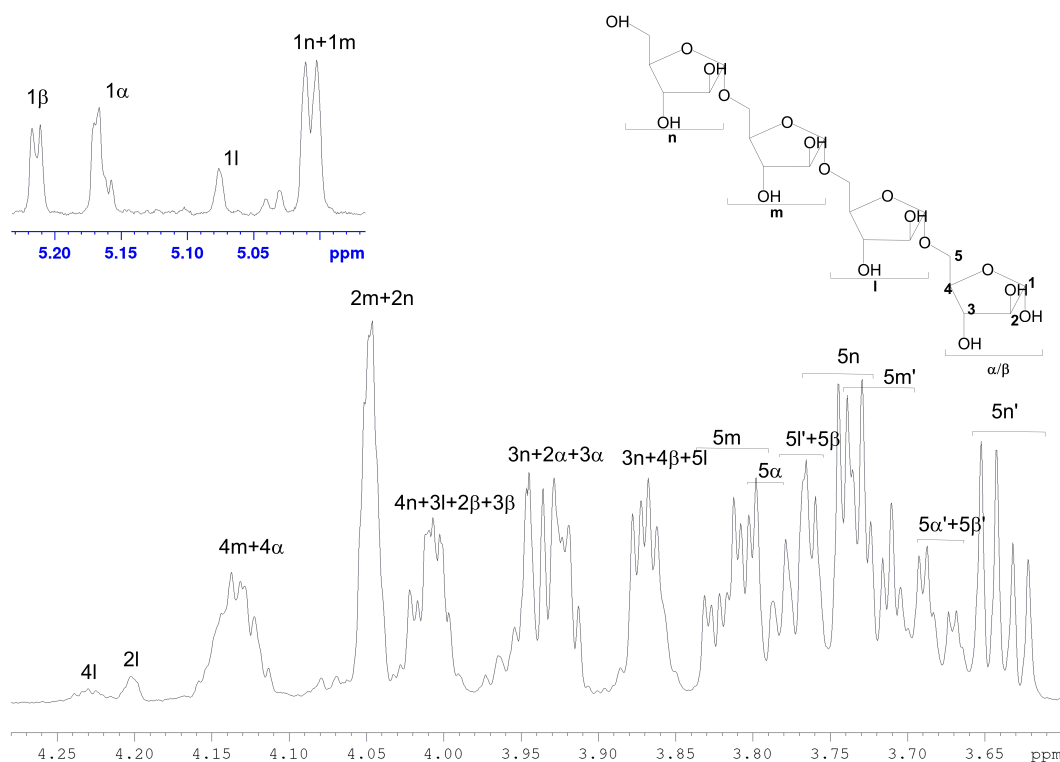


Figura II.9. Estrutura e atribuição do espectro de ^1H -RMN da arabinotetraose. O espectro foi adquirido com uma solução de 2,5 mM de arabinotetraose num espectrofotômetro de 600 MHz a 298 K com 32 scans.

Tabela II.6. Desvios químicos ^1H da arabinotetraose. Tabela com as atribuições das ressonâncias espectrais (ppm) ^1H da arabinotetraose.

Sistema de spin	H1	H2	H3	H4	H5	H5'
α	5,168	3,941	3,939	4,141	3,776	3,673
β	5,211	4,001	4,003	3,858	3,766	3,763
l	5,075	4,196	4,010	4,227	3,862	3,774
m	5,004	4,044	3,921	4,126	3,804	3,725
n	5,005	4,044	3,865	4,006	3,752	3,640

Tabela II.7. Desvios químicos ^{13}C da arabinotetraose. Tabela com as atribuições das ressonâncias espectrais (ppm) ^{13}C da arabinotetraose.

Sistema de spin	C1	C2	C3	C4	C5	C5'
α	101,138	81,440	75,930	81,210	67,000	67,000
β	95,329	76,040	74,410	79,420	68,340	68,340
l	107,050	79,120	82,320	81,560	66,450	66,450
m	107,44	80,880	76,720	82,280	66,840	66,840
n	107,44	80,880	75,930	83,900	61,150	61,150

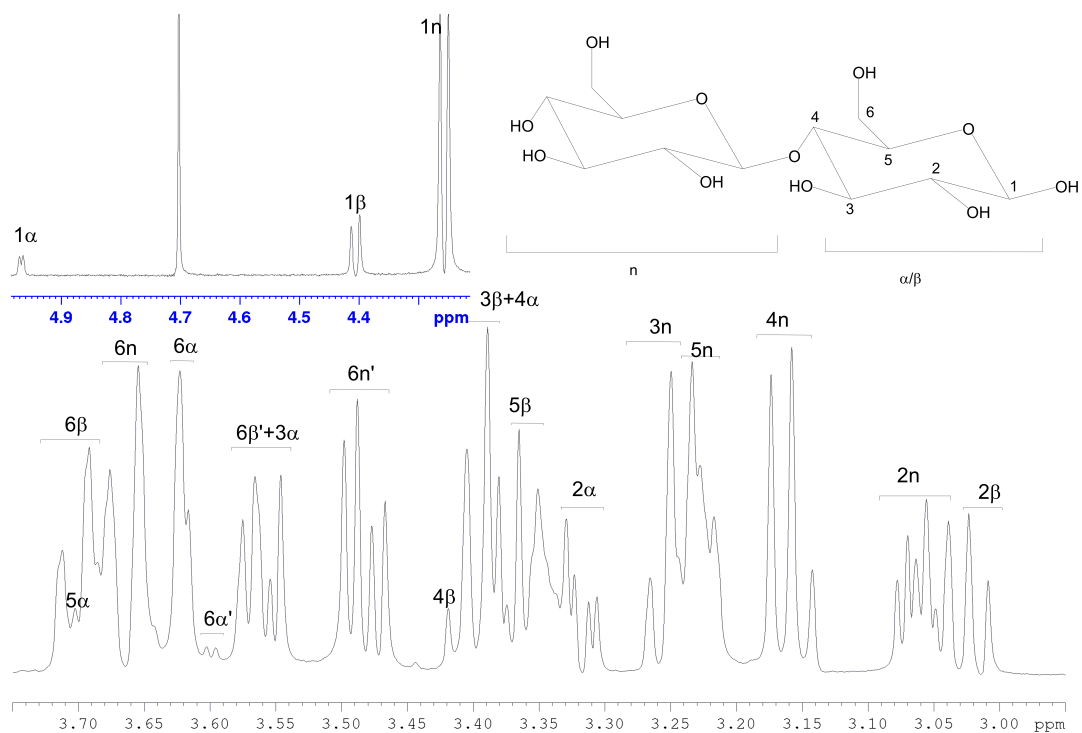


Figura II.10. Estrutura e atribuição do espectro de ^1H -RMN da celobiose. O espectro foi adquirido com uma solução de 1 mM de celobiose num espectrofotômetro de 600 MHz a 298 K com 32 scans.

Tabela II.8. Desvios químicos ^1H da celobiose. Tabela com as atribuições das ressonâncias espectrais (ppm) ^1H da celobiose.

Sistema de spin	H1	H2	H3	H4	H5	H6	H6'
α	5,20 (3,8)	3,55 (3,9, 9,8)	3,80 (9,6)	3,61	3,92	3,85	3,82 (5,3)
β	4,63	3,25 (8,6)	3,61 (9,3)	3,64	3,58	3,93 (2,2, 12,2)	3,78 (5,1, 12,3)
n	4,48 (8,6)	3,29	3,48	3,39	3,45	3,89 (12,0)	3,71 (5,9, 12,5)

Tabela II.9. Desvios químicos ^{13}C da celobiose. Tabela com as atribuições das ressonâncias espectrais (ppm) ^{13}C da celobiose.

Sistema de spin	C1	C2	C3	C4	C5	C6	C6'
α	91,80	71,21	71,32	78,62	70,09	59,85	59,85
β	95,73	73,87	74,27	78,63	74,78	60,06	60,06
n	102,53	73,16	75,49	69,46	75,97	60,56	60,56

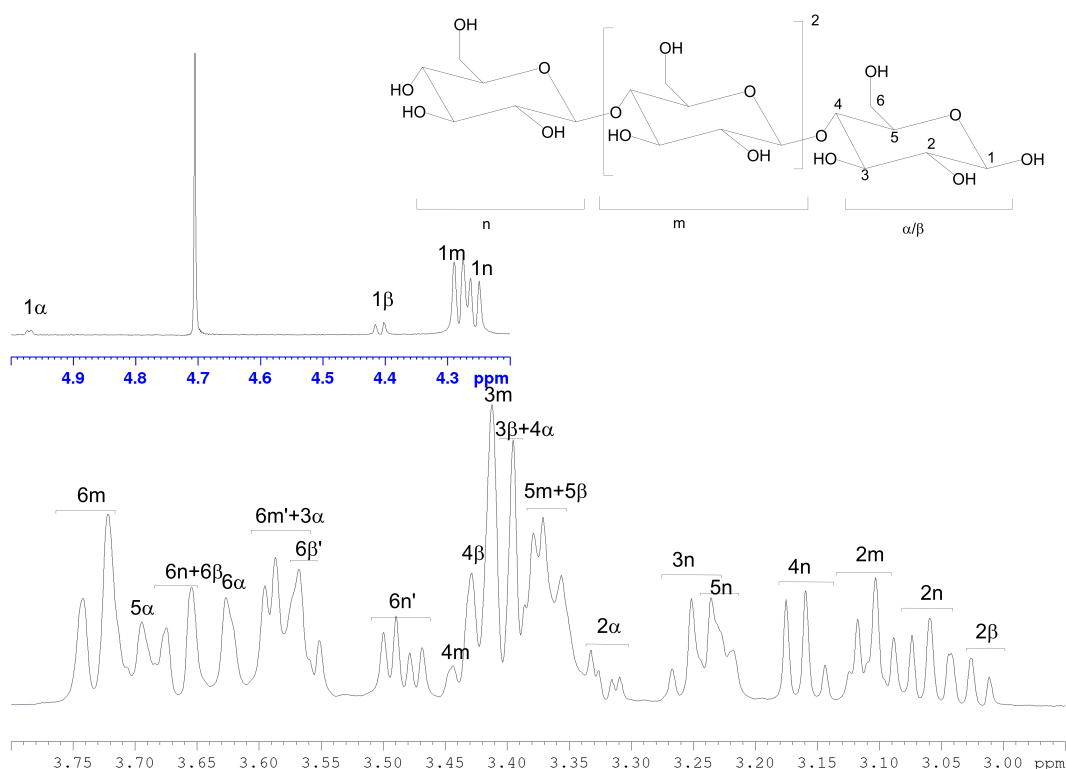


Figura II.11. Estrutura e atribuição do espectro de ^1H -RMN da celotetraose. O espectro foi adquirido com uma solução de 1 mM de celotetraose num espectrofotômetro de 600 MHz a 298 K com 32 scans.

Tabela II.10. Desvios químico ^1H da celotetraose. Tabela com as atribuições das ressonâncias espectrais (ppm) ^1H da celotetraose.

Sistema de spin	H1	H2	H3	H4	H5	H6	H6'
α	5,20 (3,8)	3,55 (3,8, 9,8)	3,80 (9,5)	3,62 (9,5)	3,92 (9,7)	3,85	3,83
β	4,63 (8,0)	3,26 (8,7)	3,61 (9,6)	3,65	3,57	3,93 (11,0)	3,78 (5,0, 12,2)
m	4,51 (8,0)	3,33 (8,6)	3,64 (8,4)	3,68	3,60	3,95 (2,0, 12,3)	3,80 (5,0, 12,4)
n	4,48 (8,0)	3,29 (8,7)	3,48 (9,1)	3,39 (9,5)	3,44	3,89 (12,3)	3,71 (5,9, 12,4)

Tabela II.11. Desvios químicos ^{13}C da celotetraose. Tabela com as atribuições das ressonâncias espectrais (ppm) ^{13}C da celotetraose.

Sistema de spin	C1	C2	C3	C4	C5	C6	C6'
α	91,72	71,17	71,34	78,29	70,04	59,85	59,85
β	95,76	73,92	74,09	78,29	74,73	59,85	59,85
m	102,24	72,95	74,09	78,29	74,73	59,85	59,85
n	102,58	73,12	75,38	69,39	75,87	60,50	60,50

II.3.3. Estudos de STD-RMN

A AraN é uma SBP e sabe-se através de estudos de mutagénese que se encontra envolvida no processo de reconhecimento molecular de arabinoligossacáridos, não reconhecendo em particular a arabinose, tendo um papel crucial na importação dos mesmos em *Bacillus subtilis* [9].

No entanto há algumas questões que ainda não estão respondidas nomeadamente: i) especificidades/promiscuidade de substratos reconhecidos (reconhecimento de pentaoses na forma de furano, bem como pentaoses e hexaoses na forma de pirano), ii) quais os epítomos do ponto de vista do ligando que se encontram a interagir com a AraN, nomeadamente se o processo de reconhecimento é efectuado através das subunidades terminais (redutora ou não redutora) ou se preferencialmente este reconhecimento é efectuado através das subunidades centrais iii) qual a natureza dessas mesmas interações, iv) qual o tamanho ótimo do oligosacáridos no processo de reconhecimento molecular, permitindo caracterizar o tamanho do local de ligação v) quão forte é a afinidade da AraN para os arabinoligossacáridos.

Para aferir se a AraN é uma SBP com uma especificidade mais ou menos alargada foram testados oligossacáridos de natureza diversa (arabinoligossacáridos, celoligossacáridos, xiloligossacáridos, manoligossacáridos e maltoligossacáridos). O mecanismo de reconhecimento é elucidado com o auxílio de GEM-Group epitope mapping, permitindo por um lado identificar os epítomos envolvidos no processo de reconhecimento molecular, e por outro quais as regiões do oligossacárido que recebem uma maior percentagem de saturação revelando que regiões monoméricas se encontram mais próximas do local de ligação. Através da análise do factor de amplificação (A_{STD}) é também possível aferir qual o tamanho ótimo do oligossacárido reconhecido pela AraN, observando-se um decréscimo deste parâmetro à medida que o carbohidrato fica mais acessível ao solvente (fora do local de ligação). A afinidade do processo de reconhecimento molecular foi aferida ao longo da titulação da AraN com arabinotriose permitindo recolher a magnitude do valor de K_d . Finalmente foram efectuados estudos de competição para validar: i) que o reconhecimento de outros carbohidratos para além de arabinoligossacáridos não é devido a interações inespecíficas, e ii) para comparar a afinidade destes outros carbohidratos relativamente a arabinoligossacáridos.

II.3.3.1. Optimização do tempo de saturação e mapeamento de epítomos

Antes de se proceder à aquisição dos espectros de STD-RMN para efeitos de cálculo de constantes de dissociação ou para o mapeamento de epítomos, existem um conjunto de parâmetros experimentais que necessitam de optimização, sendo um deles o tempo de saturação. A análise mais detalhada dos factores que afectam a transferência de saturação da proteína para o ligando é útil para entender a necessidade desta optimização.

Em solução a transferência de magnetização depende essencialmente do *turn-over* que caracteriza o processo de reconhecimento do ligando por parte da proteína, e em última análise da constante de dissociação da proteína ($K_d = k_{off}/k_{on}$). Quanto maior for o *turn-over* e consequentemente maior o k_{off} maior são as probabilidades de se acumularem em solução moléculas de ligando que não permaneceram tempo suficiente no local de ligação da proteína, e consequentemente não foram

saturadas. Por outro lado se a constante de afinidade for muito elevada, não se irão acumular moléculas de ligando saturadas em solução suficientes para reproduzir um espectro de STD-RMN. Por esta razão a técnica de STD-RMN só é reprodutível para sistemas em permuta rápida, que pode ser traduzida por processo de reconhecimento com K_d entre 10^{-3}M e 10^{-8}M [115, 116].

Para além da eficiência da transferência de magnetização, a intensidade do espectro de STD está também dependente do número de moléculas de ligando que recebem a magnetização. Assim, a sensibilidade da técnica de STD pode ser aumentada ora pela utilização de um excesso de ligando ora pela utilização de um tempo de saturação maior. A utilização de um excesso de ligando diminui a probabilidade de um ligando saturado voltar a entrar no local de ligação (*re-binding*) perdendo-se com isso amplificação da saturação de ligandos. Por outro lado, a utilização de um maior tempo de saturação permite que uma maior fração de moléculas de ligando receba saturação por parte do receptor.

Para uma experiência de STD bem sucedida estes parâmetros têm de ser otimizados. A escolha do excesso de ligando é normalmente feita tendo em conta a ordem de grandeza da K_d visto que a saturação do receptor irá levar em última análise à depleção da amplificação da saturação dos ligandos em solução. Por outro lado a escolha de um tempo de saturação muito elevado pode ter um impacto negativo nas moléculas de ligando em solução visto que estas podem perder a magnetização por processos de relaxação [117, 118].

A amplificação da magnetização é traduzida pelo factor de amplificação de STD (A_{STD}) [119], sendo este valor em média o número de moléculas de ligando saturadas por molécula de receptor, traduzido pela equação 1:

$$A_{STD} = \frac{I_0 - I_{SAT}}{I_0} \times \frac{[L]}{[P]} = \frac{I_{STD}}{I_0} \times \frac{[L]}{[P]} \quad \text{Equação 1}$$

A fração $(I_0 - I_{SAT})/I_0$, representa em última análise o efeito de STD num dado protão do ligando, isto é, esta fração é representativa da intensidade de um dado protão no espectro de STD (I_{STD}) comparativamente à mesma ressonância no espectro de referência onde o ligando não foi saturado (I_0). O mapeamento de epítomos [119] é efectuado com base neste factor. O protão com o maior A_{STD} é igualado a 100%, sendo os restantes normalizados relativamente ao sinal de maior intensidade.

Para determinar o tempo de saturação óptimo realizaram-se várias experiências de STD-RMN com diferentes tempos de saturação (secção II.2.6.2.), registando-se a amplificação da saturação do ligando em solução em função do tempo de saturação. Nesta experiência, utilizou-se um excesso 154 equivalentes de arabinotriose em relação à AraN. Na figura II.12, encontra-se representado a amplificação das ressonâncias espectrais (A_{STD}), calculadas de acordo com a equação 1 para cada ressonância, em função do tempo de saturação. As curvas foram obtidas a partir do ajuste da equação 2 [120] para o $A_{STD}(t_{sat})$ através do método dos mínimos quadrados.

$$A_{STD}(t_{sat}) = STD^{MAX} \cdot (1 - e^{(-k_{sat}t_{sat})}) \quad \text{Equação 2}$$

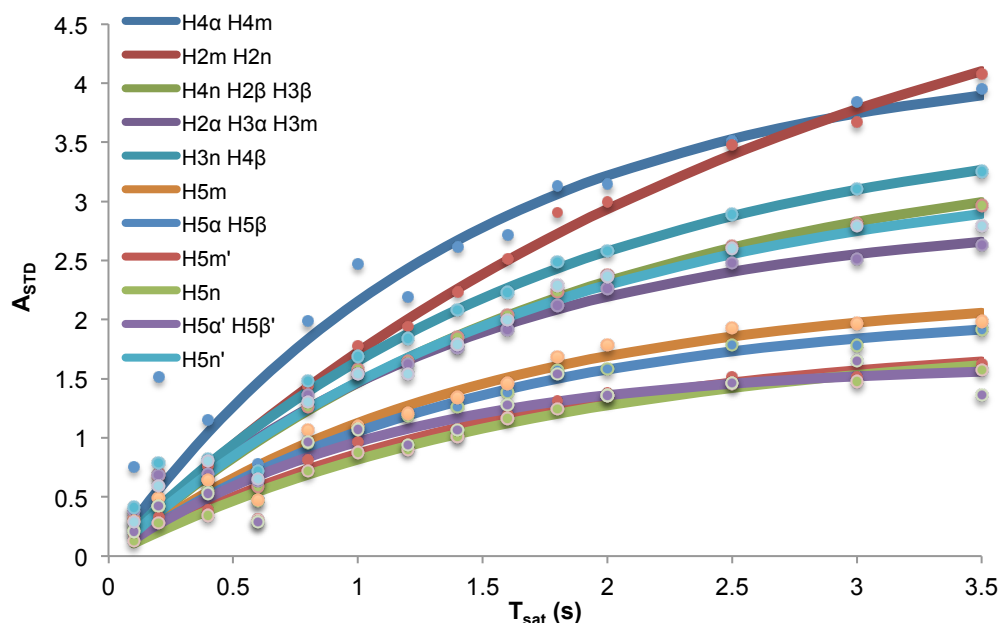


Figura II.12. Factor de amplificação (A_{STD}) em função do tempo de saturação (T_{sat}). Amplificação dos sinais dos protões do ligando em estudo em função do tempo de saturação. As curvas foram obtidas através do tratamento dos dados experimentais pela equação $(STD(t_{sat})=STD^{MAX}(1-\exp(-k_{sat}.t_{sat}))$, segundo o método dos mínimos quadrados.

Da análise do gráfico verifica-se que até ao ponto de inflexão das curvas, a amplificação dos ligandos aumenta linearmente com o aumento do tempo de saturação. Isto é, quanto maior for o tempo de saturação a que a proteína está sujeita, maior será o número de moléculas de ligando que após reconhecimento por parte do receptor recebem transferência de saturação do receptor, levando à acumulação de uma população de moléculas de ligando saturadas em solução. São estes ligandos saturados em solução que irão dar origem ao espectro de I_{SAT} . Como foi dito anteriormente, este processo de acumulação macroscópica de moléculas em solução está dependente da afinidade que o ligando tem para o receptor. Se o *turnover* do processo de reconhecimento não seguir um modelo de permuta rápida, as moléculas de ligando são saturadas, no entanto perdem essa magnetização por processos de relaxação longitudinal antes mesmo de se acumularem na solução, não dando consequentemente origem a um espectro I_{SAT} . Visto que o espectro final I_{STD} resulta da diferença do espectro I_0 e I_{SAT} , então um espectro de STD sem ressonâncias espectrais visíveis, pode refletir dois acontecimentos: i) o receptor não reconheceu o ligando e consequentemente não transferiu a magnetização para o último, ou ii) a afinidade entre o receptor e o ligando é de tal modo elevada que a constante temporal k_{off} é inferior á taxa de relaxação longitudinal (R_1).

Assim como é possível observar na figura II.12 em cima representada, para tempos de saturação superiores a 3 segundos atinge-se um *plateau*, indicativo que as moléculas que receberam saturação por parte do receptor num instante $t=0s$ terão relaxado ao final de um tempo $t=3s$, isto é, à medida que uns ligandos são saturados outros relaxam, atingindo-se um tipo de estado estacionário.

Para os estudos posteriores foi então escolhido um tempo de saturação de 1,5 segundos, visto que representa o compromisso em que para um máximo de A_{STD} , todos os protões do ligando estão em condições de velocidade inicial (minimização dos efeitos de relaxação) observando-se uma discriminação do factor A_{STD} para todos eles. A discriminação do valor de A_{STD} é extremamente

importante para o mapeamento dos epítomos e como tal vai ser objecto de uma análise mais aprofundada (II.3.3.4).

Como é possível observar na figura II.13 existe uma diferenciação da amplificação da saturação em solução para os diferentes protões do ligando. Através do ajuste da equação 2 aos pontos experimentais foram obtidos dois parâmetros: i) STD^{MAX} , que representa a amplificação máxima individual de cada protão e ii) k_{sat} que representa a velocidade (taxa) com que ocorre o *build-up* em solução, da saturação do protão do ligando à medida que o tempo de saturação aumenta [121]. Este valor pode ser retirado da equação 3.

$$\frac{\partial A_{STD}}{\partial t(0)} = \frac{\partial}{\partial t} (STD^{MAX} - STD^{MAX} \cdot e^{-(k_{sat}t_{sat})}) = STD^{MAX} k_{sat} \cdot e^{-(k_{sat}t_{sat})} = STD^{MAX} k_{sat} \quad \text{Equação 3}$$

Diferentes valores de k_{sat} para diferentes protões do ligando podem estar associados a diferentes eficiências de transferência de magnetização do receptor, sendo ainda que este valor não é modulado pelo T_1 específico de cada protão. A recta tangente à isotérmica de ligação de um dado protão para um $t_{sat}=0$ irá em última análise resultar do produto $k_{sat} \cdot STD^{MAX}$ - parâmetro STD_{fit} . Na figura II.13 é possível observar este parâmetro para cada conjuntos de protões (na secção VI.4, encontra-se descrito o A_{STD} , k_{sat} , STD^{MAX} , STD_{fit} em função de cada t_{sat} , para cada conjunto de protões).

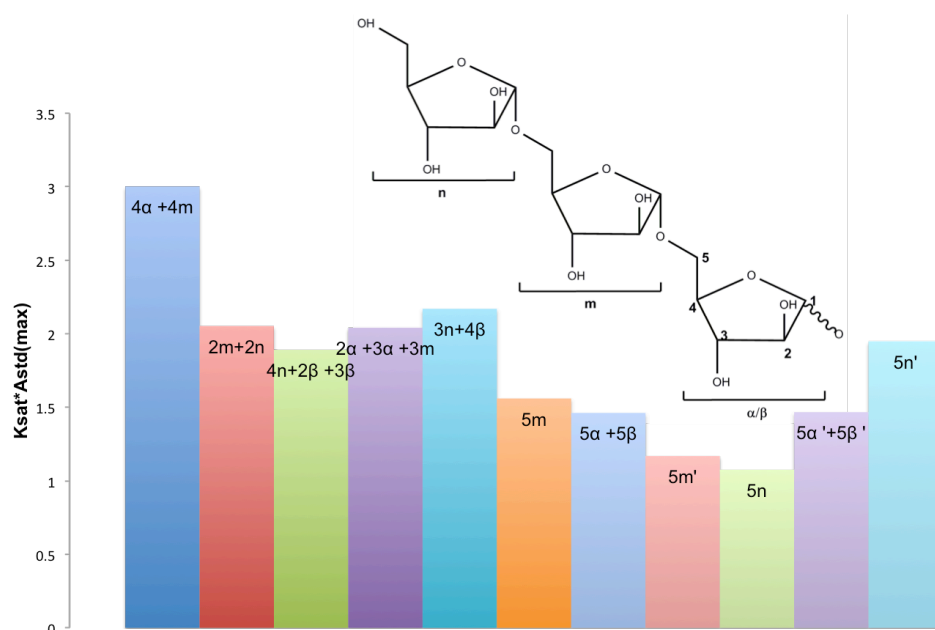


Figura II.13. STD_{fit} de cada protão. Gráfico de barras representativo do STD_{fit} para cada protão da arabinotriose em estudo.

Como vimos anteriormente, à medida que o tempo de saturação aumenta maior será a A_{STD} , no entanto maiores serão por um lado os efeitos de T_1 , e por outro lado maior será a probabilidade de ocorrer um processo de re-ligação de uma molécula de ligando previamente saturada, não recebendo esta última a mesma saturação que uma molécula de ligando não saturada, fenómeno este que contribui com uma quebra na amplificação de ligando saturado em solução e uma subestimação do valor de A_{STD} . Idealmente, um correto mapeamento de epítomos de ligação deveria ser efectuado com

uma correção do A_{STD} relativamente ao T_1 de cada protão, ou alternativamente o tempo de saturação utilizado deveria ser inferior ao processo de relaxação longitudinal dos núcleos envolvidos em estudo. Esta última hipótese acarreta algumas desvantagens, quanto menor for o tempo de saturação, menor será a amplificação de moléculas de ligando saturadas em solução, e consequentemente o espectro de STD apresentaria uma menor razão sinal ruído, o que invariavelmente iria introduzir erros durante o GEM [118].

O $STD_{fit}(k_{sat} * STD^{MAX})$ contorna estes problemas pois é obtido como a equação 3 especifica, num instante $t_{sat}=0$, em condições de velocidade inicial onde processos de re-ligação ao local de ligação e processos de relaxação T_1 não surtem efeitos na obtenção de A_{STD} .

O mapeamento de epítomos, conforme se encontra na tabela II.12, é efectuado através da razão entre as intensidades do espectro STD ($I_{STD}=I_0-I_{SAT}$) e do espectro de referencia I_0 , sendo que o protão ou conjunto de protões que evidenciarem um maior A_{STD} é normalizado para 100%, e os restantes são normalizados relativamente a este.

Tabela II.12. Mapeamento de epítomos para diferentes T_{sat} . Tabela com o mapeamento de epítomos (normalizado em %), recorrendo a valores de A_{STD} para um limite de tempo de saturação zero (STD_{fit}), e para diferentes tempos de saturação.

Epítopo	$T_{sat}=0,6s$ (%)	$T_{sat}=2s$ (%)	$T_{sat}=3,5s$ (%)	STD_{fit} (%)	Taxa (%)
H4 α +H4m	100,0	100,0	96,9	100,0	3,2
H2m+H2n	82,1	95,3	100,0	68,4	-17,9
H4n+H2 β +H3 β	75,6	75,5	72,8	63,0	3,9
H2 α +H3 α +H3m	81,6	72,0	64,6	68,0	26,3
H3n+H4 β	92,3	82,2	79,8	72,3	15,7
H5m	60,9	56,8	48,7	51,9	24,9
H5 α +H5 β	60,1	50,5	47,0	48,7	27,9
H5m'	39,9	43,8	39,8	38,9	0,1
H5n	39,6	43,2	38,6	35,8	2,5
H5 α' +H5 β'	37,2	43,3	33,4	48,8	11,2
H5n'	84,2	75,2	68,5	64,9	22,9

Pela análise da tabela apresentada anteriormente, é possível observar que o mapeamento pode ser completamente diferente em função do tempo de saturação escolhido ($t_{sat}=0,6$ e $t_{sat}=3,5$). Em condições de $t_{sat}=0s$ (coluna de STD_{fit}) os protões H4 α e H4m, são aqueles que estarão em maior proximidade da proteína, no entanto quando a mesma medição é efectuada com um tempo de saturação de 3,5s, há um claro decaimento do factor de amplificação evidenciando os processos de relaxação T_1 , referidos anteriormente, subestimando a sua interação e proximidade com a proteína. Assim pode-se concluir que o t_{sat} quando muito elevado pode ser fortemente modelado por efeitos de T_1 , sendo que neste caso certamente os protões H2m e H2n terão um tempo de relaxação longitudinal mais elevado que os protões H4 α e H4m. A tendência é a de que protões com menores T_1 aumentem a sua intensidade relativa A_{STD} , quando são usados tempos de saturação mais curtos, ao passo que protões com T_1 mais elevados vejam as intensidades A_{STD} reduzirem.

Assim pela análise da tabela II.12 é possível constatar que há um aumento de 25% e 28% do valor de A_{STD} quando da diminuição do t_{sat} de 3,5s para 0,6s, para os prótons H5m bem como para os prótons H5 α e H5 β respectivamente. Por seu turno, os prótons H2m e H2n registam uma queda de A_{STD} de 18%. Assim, e visto que este ensaio de AraN foi efectuado com um excesso de 154 de arabinotriose, podemos desde já eliminar como factor das variações do A_{STD} processos de religação, concluindo que os processos que se encontram a modelar o A_{STD} resultam das diferenças de relaxação longitudinal intrínseca de cada núcleo. Em suma, pode-se prever que a ordem crescente de T_1 (s) H5 α +H5 β <H5m<H2m+H2n. Através da análise da figura II.12, é possível constatar que o único conjunto de prótons que não atingem um *plateau* após um tempo de saturação de 3,5 segundo são os prótons H2n e H2m, o que pode corroborar o facto destes prótons terem tempos de relaxação longitudinal superiores aos restantes prótons do arabinotriose, o que em última análise poderá sobrestimar a sua interação com a AraN. Em conclusão quanto menor for T_1 mais rápido o próton atinge o *plateau* [118].

A análise detalhada e a interpretação do epítipo obtido para a arabinotriose no contexto da interação com a AraN será apresentada posteriormente na secção II.3.3.4 e II.3.3.5.

II.3.3.2. Influência da temperatura

A influência da temperatura de aquisição dos espectros de STD-RMN no efeito de STD detectado foi outro dos parâmetros estudado. A cinética de reconhecimento, bem como a constante de afinidade da formação do complexo receptor-ligando é fortemente influenciada pela temperatura, e consequentemente reflete-se no efeito de STD observado. Visto que a técnica de STD assenta na transferência de saturação do receptor para o ligando, é fulcral que a escala de tempo deste processo seja otimizada, podendo-o ser através da variação de temperatura. Um exemplo claro é o da interação fraca entre o receptor e o ligando, que pode levar em última análise à incapacidade do ligando receber saturação, não se registando efeito de STD. Nesta situação a diminuição da temperatura pode ser benéfica na medida em que irá alterar a dinâmica do processo de reconhecimento de modo a diminuir o valor da constante k_{off} , permitindo ao ligando um maior tempo de residência em contacto com o receptor e como tal receber a saturação de um modo mais eficiente [122]. O inverso para o caso de um ligando com uma elevada afinidade também é verdade. Neste caso, a saturação recebida na forma ligada (PL), irá amplificar-se em solução se a temperatura for aumentada.

Na figura II.14, são apresentados os espectros de STD-RMN obtidos para a arabinotriose às temperaturas de 35°, 25° e 10° C.

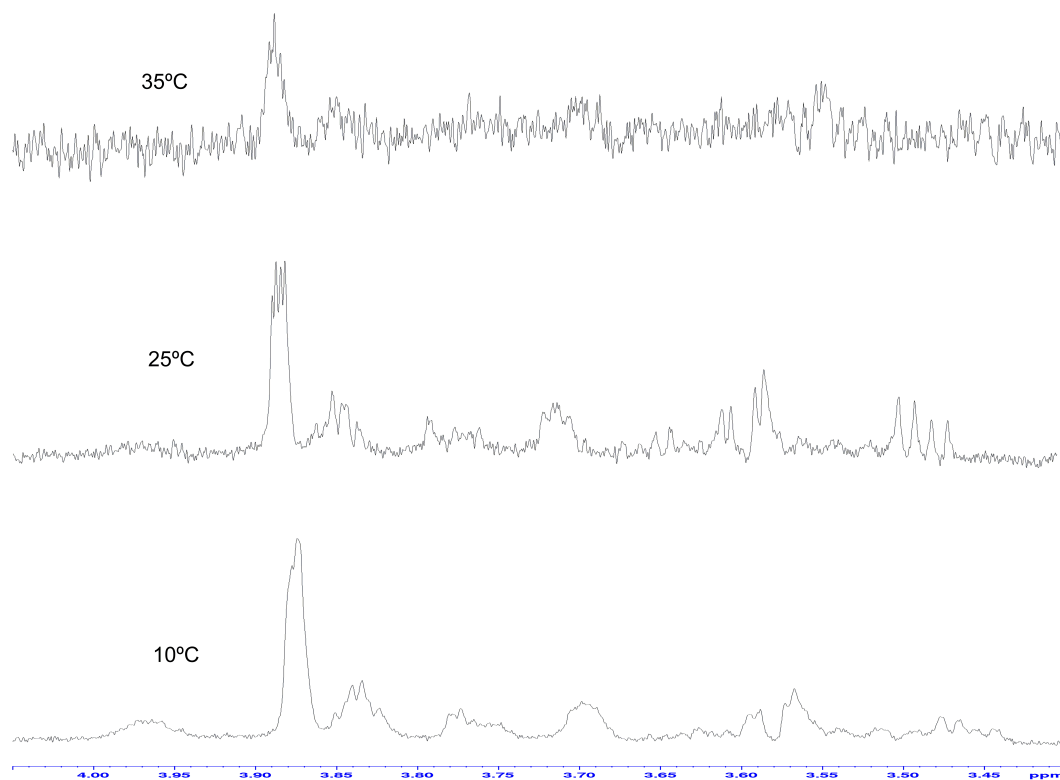


Figura II.14. Efeito da temperatura no espectro de STD-RMN. Espectros de ^1H -STD-RMN obtidos para a arabinotriose (2,5 mM) na presença de AraN (25 μM) a diferentes temperaturas. Os espectros ^1H -STD (I_{STD}) foram resultado da subtração do espectro ^1H -STD *off-resonance* (I_0), com irradiação a 20000Hz e do espectro *on-resonance* ^1H -STD (I_{SAT}) com obtido pela irradiação na região espectral a 150Hz.

A análise teórica do comportamento observado na figura II.14, pode ser explicada através da relaxação longitudinal *spin-lattice* T_1 bem como através da relaxação transversal *spin-spin* T_2 , através das quais o equilíbrio termodinâmico é alcançado após uma perturbação inicial, sob a forma de um pulso de radiofrequências. A relaxação T_1 está associada ao re-establishimento da distribuição das populações de spins nucleares ao nível do equilíbrio térmico de Boltzmann, ao passo que a relaxação T_2 está correlacionada com a perda de coerência dos spins nucleares ao nível da magnetização transversal (consultar secção, VI.1.2. *Relaxação longitudinal T_1 e relaxação transversal T_2*).

O principal mecanismo através do qual ocorre relaxação em NMR é através da presença de um campo magnético oscilatório à frequência de Larmor promovendo consequentemente transições entre os níveis energéticos das populações α e β dos spins nucleares (consultar secção VI.1.4. *Movimento molecular e relaxação de spin* e VI.1.5. *Tempo de correlação e densidade espectral*).

Através da análise da figura II.15 podemos tirar conclusões acerca do movimento molecular rotacional e translacional da arabinotriose em função da variação da temperatura e viscosidade. Moléculas de pequenas dimensões apresentam um movimento molecular rápido (τ_C pequeno), ao passo que moléculas de elevado peso molecular como são o caso das proteínas apresentam um movimento molecular lento (τ_C elevado). Associado a este tempo de correlação (τ_C) está a capacidade de se gerarem campos magnéticos oscilatórios capazes de inferirem transições energéticas ao nível das distribuições populacionais de spins nucleares.

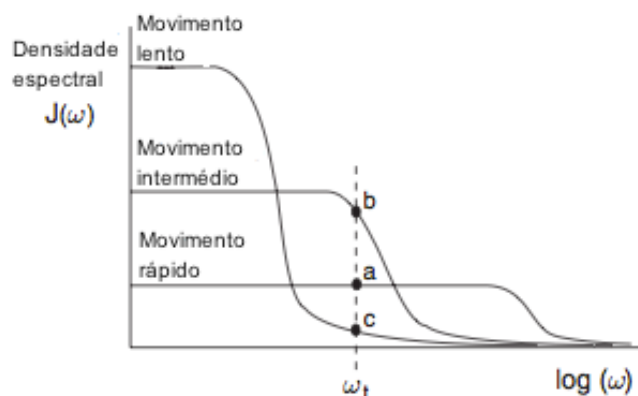


Figura II.15. Variação da densidade espectral com o movimento molecular em função da frequência de Larmor. Representação esquemática da densidade espectral em função da frequência de oscilação dos campos magnéticos, para diferentes regimes de movimento molecular, com ω_L a representar a frequência de Larmor.

Moléculas pequena com um movimento molecular rápido apresentam uma pequena componente de campos oscilatórios à frequência de Larmor (ω_L), capazes de induzir relaxação por T_1 (região a da figura II.15). À medida que o movimento molecular diminui a densidade espectral começa por aumentar em ω_L num regime de movimento molecular intermédio proporcionando consequentemente a diminuição T_1 levando a uma relaxação de spins nucleares mais rápida. Porém se o regime de movimento molecular continuar em decréscimo a frequência de campos magnéticos oscilatórios em ressonância com a frequência de Larmor irá descer drasticamente, levando a um novo aumento de T_1 com um aumento de tempo de relaxação. Este comportamento oscilatório de T_1 pode ser observado na figura II.16 na curva a cheio.

O mecanismo de relaxação T_1 , tal como em T_2 é modelado pela temperatura e viscosidade visto que estes factores modelam o movimento molecular em solução: i) o aumento da temperatura e/ou diminuição da viscosidade do meio irá promover o *tumbling* molecular em solução, e consequentemente irá diminuir o τ_C de um modo artificial.

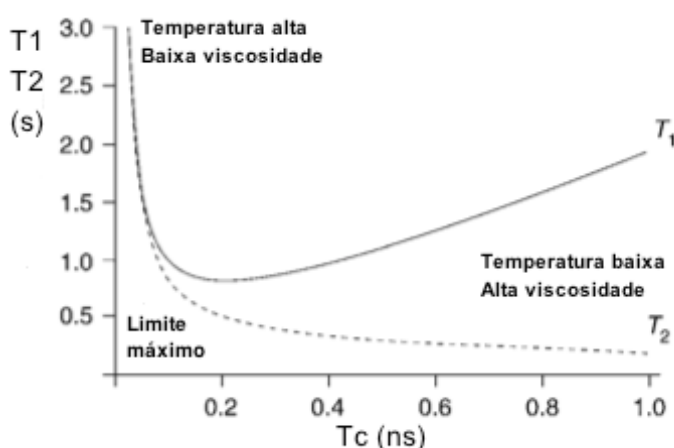


Figura II.16. Efeito do tempo de correlação no tempo de relaxação T_1 e T_2 . Comportamento de T_1 e T_2 em função do tempo de correlação (τ_C), este ultimo modelado pela temperatura e viscosidade da amostra.

Através da análise da figura II.14, é possível constatar que a 10°C à um alargamento da linha, resultante do aumento do τ_C da arabinotriose em solução, algo que também poderia ocorrer

através do aumento da viscosidade, caso a concentração de proteína fosse muito elevado por exemplo.

O facto do espectro a 37°C apresentar uma baixa razão sinal/ruído está associado ao facto do processo de ligação de reconhecimento por parte da AraN ficar mais “solto” sendo que a arabinotriose não é eficientemente saturada pelo receptor, visto que não fica o tempo suficiente no local de ligação. A arabinotriose é uma molécula de relativamente pequena encontrando-se na região *extreme narrowing* da figura II.16, assim um aumento de temperatura e consequente diminuição do tempo de correlação apenas iria levar ao aumento de T_1 e T_2 o que me leva a concluir que a baixa razão sinal/ruído presente no espectro da figura II.14, está associada à natureza do processo de reconhecimento e não a factores de relaxação.

Este resultado, associado aos baixos valores de A_{STD} determinados (secção II.3.3.1 - A_{STD} não atinge sequer o valor 5 para o tempo de saturação mais alto) são sinónimos de uma interação relativamente fraca entre a arabinotriose e a AraN.

Tabela II.13. Factor de amplificação (A_{STD}) em função da temperatura. Para cada temperatura (10°C, 25°C e 37°C) são apresentados os valores de A_{STD} com e sem normalização, para cada conjuntos de epítomos.

	H4 α + H4m	H2m +H2n	H2 β +H3 β +H4n	H2 α +H3 α +H3m	H4 β +H 3n	H5m	H5 α +H 5 β	H5n	H5'm	H5 α '+ H5 β '	H5n'
10°C	2,7 59,0	4,5 100,0	2,7 60,0	1,6 34,9	2,8 61,8	1,2 27,0	1,5 32,7	1,7 37,9	1,7 36,8	1,3 27,8	2,8 62,7
25°C	1,3 75,6	1,8 100,0	1,2 64,7	0,8 43,2	1,3 73,1	0,7 38,1	0,7 39,1	0,7 41,5	0,7 40,2	0,6 32,7	1,3 73,1
37°C	0,8 100,0	0,6 76,9	0,6 79,2	0,4 55,1	0,6 72,0	0,4 56,7	0,4 50,9	0,3 39,8	0,3 36,5	0,3 38,3	0,5 65,0

Verifica-se assim que a temperatura mais indicada para obtenção dos STD's seria entre os 10°C e os 25°C. A 10°C a principal vantagem desta ser a temperatura que permite obter um maior A_{STD} , e consequentemente um menor erro na determinação do mapeamento de epítomos e K_d devido à maior razão sinal ruído, no entanto o mapeamento de epítomos poderá estar sujeito a erros devido a fenómenos de relaxação diferencial entre os vários protões. É de notar que à temperatura de 37°C o conjunto de protões com maior grau de saturação são os protões H4 α e H4m, ao passo que os protões H2m e H2n apresentam um maior grau de saturação a temperaturas de 10°C e 25°C. Hipoteticamente a temperatura pode ter influência no modo de ligação da arabinotriose à AraN.

II.3.3.3. Calculo de K_d

Na secção da optimização do tempo de saturação (II.3.3.1), observou-se que ocorre um *build-up* do factor de amplificação com o aumento do tempo de saturação. Este *build-up* também é observado ao longo de uma titulação da proteína com o aumento da concentração de ligando. Como foi dito anteriormente, o valor de A_{STD} pode ser interpretado como o número de moléculas de ligando saturadas por molécula de receptor, assim para um determinado tempo de saturação, em condição de velocidade inicial, à medida que a concentração de ligando aumenta o A_{STD} irá aumentar até atingir um máximo de amplificação (α_{STD}). O aumento de A_{STD} em função do aumento de moléculas

de ligando em solução pode estar relacionado com dois factores: i) à medida que se aumenta $[L]$, haverá mais moléculas de ligando disponíveis para receber saturação por parte do receptor e ii) há uma menor probabilidade de ligandos saturados num instante t e com uma menor capacidade de receberem saturação por parte da proteína, visto que já haviam recebido essa saturação instantes antes, voltarem a entrar no local de ligação da proteína no instante $t+\tau$.

Eventualmente, quando a $[L] \gg K_D$ a proteína atinge a capacidade máxima de *turn-over*, ficando o local de ligação saturado e a curva de titulação alcança o valor de A_{STD} máximo correspondente a α_{STD} . Para além deste ponto, a proteína está no seu máximo de *turn-over*, não conseguindo consequentemente aumentar a resposta do efeito de STD sobre a forma de I_{STD} , no entanto como o espectro *off-resonance* I_0 é proporcional a $[L]$ este irá aumentar, e o resultado é um decréscimo de A_{STD} para condições de $[L] > [L]_T$, verificando-se uma diminuição de A_{STD} .

No figura II.17 estão representados os resultados da titulação de STD-RMN da AraN com arabinotriose (de acordo com II.2.6.2).

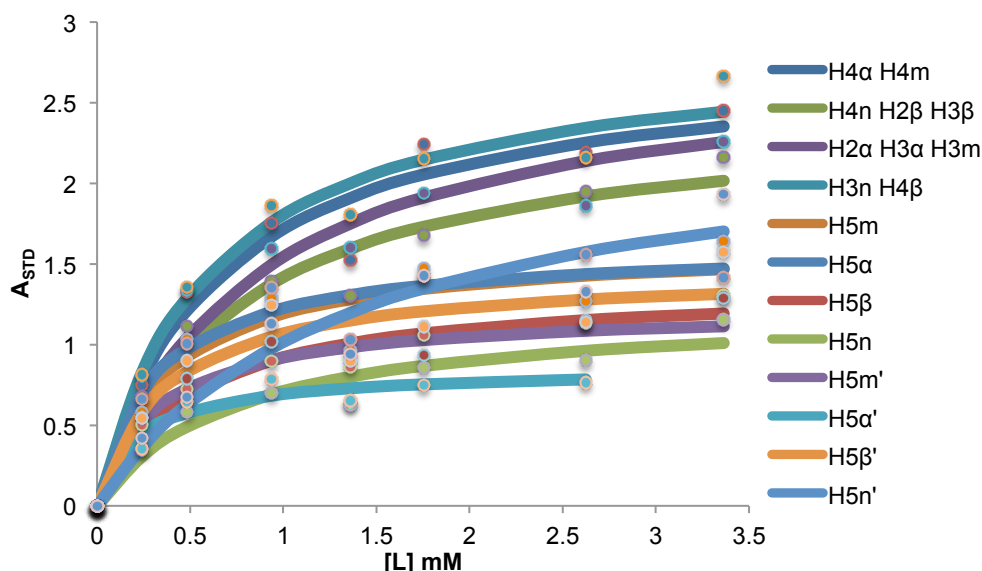


Figura II.17. Factor de amplificação (A_{STD}) em função da concentração de ligando ($[L]$). As curvas foram obtidas através do ajuste não linear dos pontos experimentais à equação $A_{STD} = (\alpha_{STD}[L]) / (K_D + [L])$, tendo os parâmetros α_{STD} e K_D sido obtidos através da minimização dos mínimos quadrados.

Através do ajuste dos valores experimentais à expressão $A_{STD} = (\alpha_{STD}[L]) / (K_D + [L])$ pode estimar-se o valor da constante de afinidade, verificando-se que se encontra numa gama de 200 μM a 700 μM . Este valor é bastante diferente do reportado na literatura que se situa entre 0,1 μM a 10 μM para o transportador de maltose MBP (*Maltose binding protein*) em *Escherichia coli*. [107].

O cálculo do valor de K_D pode ter um erro inerente associado ao facto de na fase de crescimento linear na qual não há um excesso de ligando, poderem ser observados processos de re-ligação da arabinotriose à AraN. Quando isto acontece as moléculas de arabinotriose podem ainda encontrar-se saturadas, não recebem proporcionalmente a saturação devida. Assim, existe uma subestimação do valor de A_{STD} , ou seja a amplificação do ligando saturado em solução é menor do

que aquela que deveria ser na fase inicial da titulação. Como o A_{STD} é subestimado o valor de K_d acaba por ser sobrestimado face ao seu valor real [117].

II.3.3.4. Estudo das interações ligando-AraN com arabinoligossacáridos

II.3.3.4.1. Interação com arabinose

Na figura II.18 é apresentado o espectro de STD-RMN obtido para a arabinose na presença de AraN. Podemos constatar que a arabinose na presença de AraN, dá origem a um espectro de STD sem sinais, ou seja, o espectro *off-resonance* (I_0) e o espectro *on-resonance* (I_{SAT}) são exatamente iguais pelo que a sua diferença produz um espectro I_{STD} sem nenhum sinal, o que se pode interpretar como nenhum núcleo se encontrar a interagir com o receptor molecular.

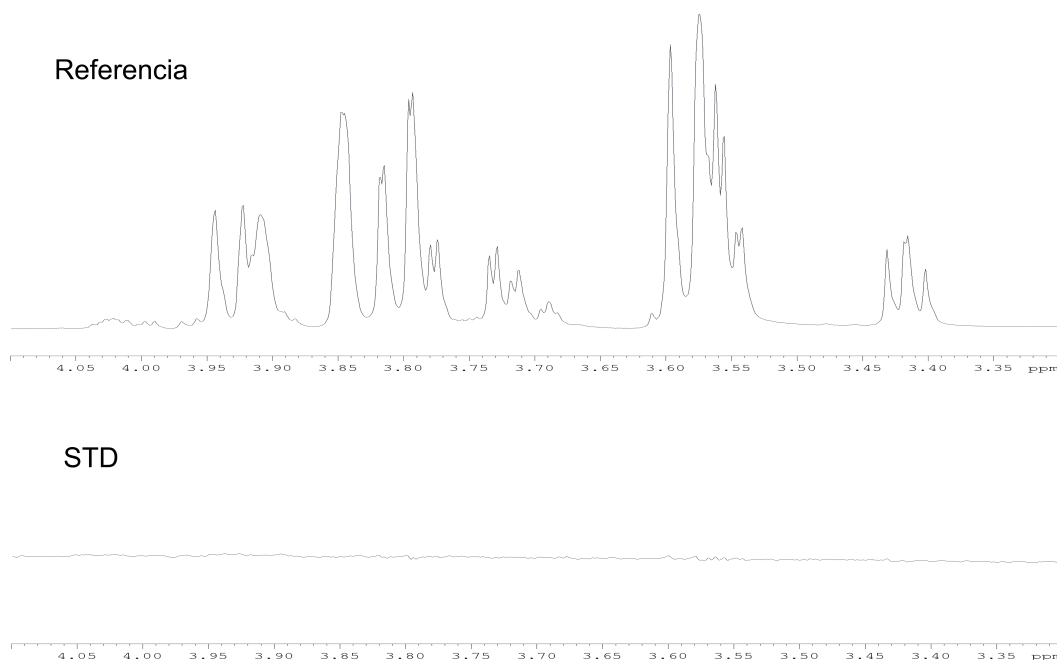


Figura II.18. Espectro de STD-RMN da arabinose com AraN. Na parte superior da figura encontra-se o espectro de referência ^1H -STD *off-resonance* (I_0), da arabinose, com irradiação a 20000Hz. Em baixo, encontra-se o espectro ^1H -STD (I_{STD}), da arabinose, resultante da diferença entre o espectro I_0 e I_{SAT} , tendo este último sido obtido pela irradiação na região espectral a 150Hz. A aquisição do espectro foi efectuada a 10°C com 2,5 mM de arabinose e 25 μM de AraN.

Este resultado está de acordo com resultados experimentais já reportados que indicam que este açúcar na sua forma monomérica não é transportado pelo sistema de transporte do tipo ABC [9], constituído pelas proteínas AraNPQ-MsmX, mas sim pelo transportador secundário AraE. Assim, é natural que este carbohidrato não seja reconhecido pela AraN, visto que esta estaria associada ao reconhecimento de substratos do transportador AraNPQ-MsmX. Neste aspecto, a AraN releva-se semelhante à proteína de reconhecimento de malto-oligosacáridos (MBP), na medida que esta reconhece maltoligossacáridos, não reconhecendo a unidade monomérica que os compõem, a D-Glucose [107].

II.3.3.4.2. Interação com arabinobiose

Na figura II.19 é apresentado o espectro de STD-RMN obtido para a arabinobiose na presença de AraN.

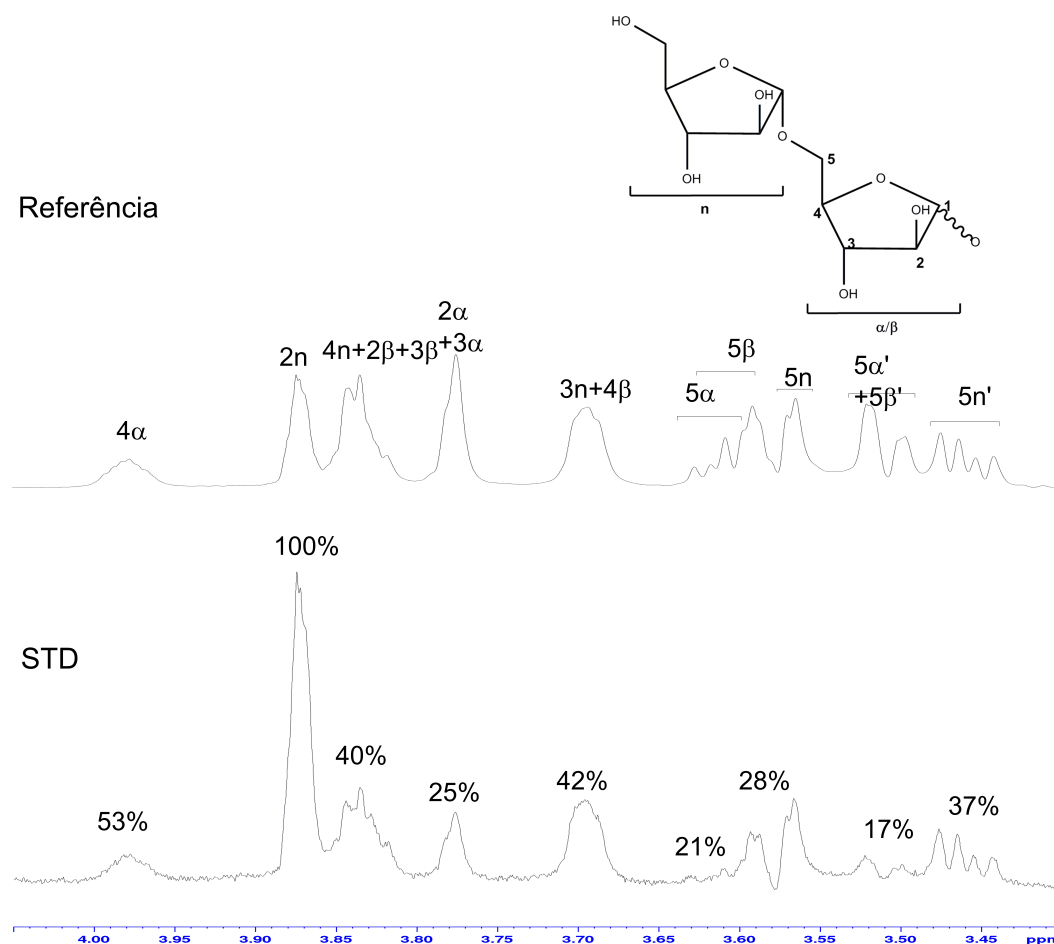


Figura II.19. Espectro de STD-RMN da arabinobiose com AraN. Na parte superior da figura encontra-se o espectro de referência ^1H -STD *off-resonance* (I_0), da arabinobiose, com irradiação a 20000Hz. Em baixo, encontra-se o espectro ^1H -STD (I_{STD}), da arabinobiose, resultante da diferença entre o espectro I_0 e I_{SAT} , tendo este último sido obtido pela irradiação na região espectral a 150Hz. A aquisição do espectro foi efectuada a 10°C com 2,5 mM de arabinobiose e 25 μM de AraN.

Relativamente à arabinobiose, esta por seu turno evidência um espectro STD que é indicativo de reconhecimento da mesma por parte da AraN. De um modo geral todas as ressonâncias espectrais presentes no espectro de STD apresentam um grau de saturação o que pode significar que todo o açúcar apresenta contribuições no processo de reconhecimento molecular pese embora, muitas das ressonâncias espectrais correspondam à sobreposição de sinais de vários protões. Esta sobreposição de sinais de protões nas mesmas regiões espectrais tem de ser tida em conta, pois na interpretação do espectro STD, a saturação responsável pelo mesmo pode ser proveniente de mais do que um protão sem que nunca se saiba qual dos protões em questão apresenta o maior efeito.

No espectro de STD da arabinobiose, é possível observar que o protão que recebe maior saturação por parte da AraN durante o processo de reconhecimento molecular é o H2n da subunidade não-redutora. Os protões dos metilenos (H5 e H5') que apresentam um elevado efeito STD pertencem também eles à subunidade não-redutora (H5n= 28% e H5n'=37%). Este dados vão

de encontro ao modelo estudado na proteína de ligação a maltoligosacáridos MBP, no qual foi evidenciado que a subunidade não redutora da maltobiose, é aquela que apresenta maior número de interações com o receptor, pese embora o local de ligação da MBP seja mais profundo no local de reconhecimento das subunidades redutoras, encontrando-se esta subunidade a estabelecer um conjunto de interações por pontes de hidrogénio [107, 123]. Os resultados da interação da AraN com a arabinobiose demonstram que a subunidade redutora deste carboidrato, não irá ficar em princípio mais enterrada no interior da AraN que os prótons da subunidade n, visto que se assim fosse seria espectável observar uma maior eficiência de transferência de saturação para a subunidade redutora.

II.3.3.4.3. Interação com arabinotriose

Na figura II.20 é apresentado o espectro de STD-RMN obtido para a arabinotriose na presença de AraN. Para a arabinotriose, tal como a arabinobiose, também se verifica a presença de sinais no espectro de STD, com a particularidade de apresentar A_{STD} maiores comparativamente à arabinobiose. Estes dados estão em concordância com o reportado por *Quioco et al* [107] e *Boraston et al* [54], que verificaram que para transportadores de maltoligosacáridos, a maltotriose é o açúcar que maior afinidade apresenta com o receptor molecular. A razão pela qual isto acontece deve-se ao facto de com o aumento do tamanho do carboidrato, nomeadamente no caso de carboidratos com mais de três subunidades, a região não-redutora ficar mais exposta ao solvente pela sua exclusão do local de ligação. Assim à medida que se verifica um aumento do tamanho do oligossacárido com mais de três subunidades o valor de A_{STD} decresce. Por outro lado, devido ao facto de a arabinotriose ter mais uma subunidade que a arabinobiose, esta fica mais estabilizada no local de ligação recebendo maior saturação por parte da AraN.

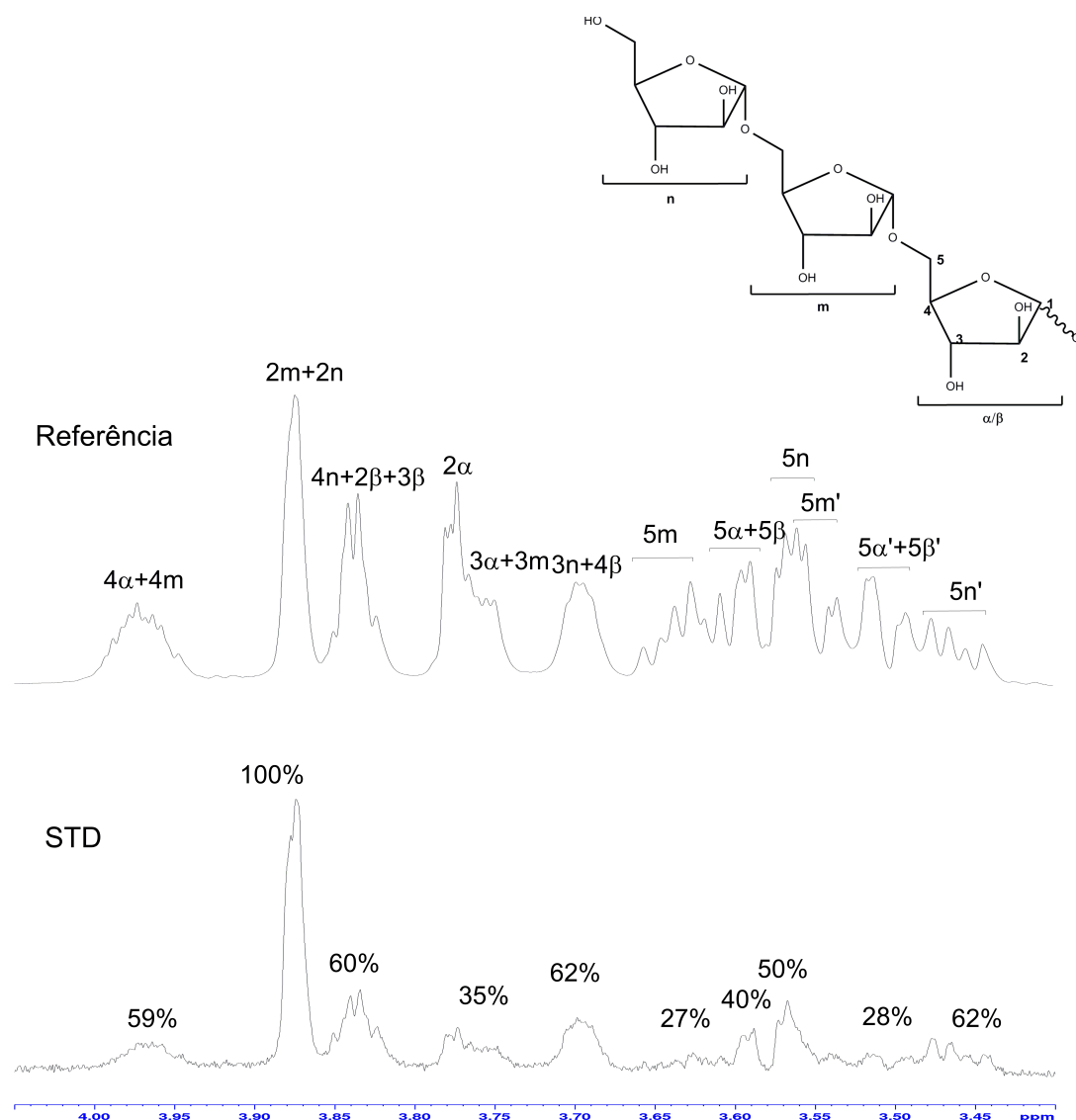


Figura II.20. Espectro de STD-RMN da arabinotriose com AraN. Na parte superior da figura encontra-se o espectro de referência ^1H -STD off-resonance (I_0), da arabinotriose, com irradiação a 20000Hz. Em baixo, encontra-se o espectro ^1H -STD (I_{STD}), da arabinotriose, resultante da diferença entre o espectro I_0 e I_{SAT} , tendo este último sido obtido pela irradiação na região espectral a 150Hz. A aquisição do espectro foi efectuada a 10°C com 2,5 mM de arabinotriose e 25 μM de AraN.

Assim como se encontra ilustrado na figura II.20, os prótons que apresentam o maior grau de saturação são os H2m e H2n, por outro lado os grupos metileno, com um maior grau de saturação são aqueles que se encontram presentes na subunidade não redutora (H5n e H5n'), com um grau de saturação de 50% e 62%. Se fosse de esperar da AraN um comportamento igual ao verificado para outras SBP cujas regiões de interação do oligosacáridos foram estudadas por cristalografia de raio-x, então seria de esperar um maior STD nos monómeros da subunidade redutora (prótons α e β) da arabinotriose, algo que vai ao encontro com os dados da secção II.3.3.1 (tabela II.12).

Como esperado, face à natureza da interação dos carboidratos com proteínas do tipo das SBP, ou seja, interações muito fortes devido a um consórcio de pontes de hidrogénio (associadas ao lobo N-terminal), bem como de interações de Van der Waals (associadas ao lobo C-terminal), todo o ligando irá estar envolvido num complexo mapa de interações. No caso da maltotriose da MBP, este ligando encontra-se envolvido em 21 pontes de hidrogénio (cooperativas, e grande parte delas

mediadas por moléculas de água), envolvidas no reconhecimento da subunidade redutora, bem como 80 interações Van der Waals, envolvidas no reconhecimento da subunidade central e da terceira subunidade a partir da subunidade redutora [107].

Do mesmo modo, para a AraN, na série arabinose, arabinobiose e arabinotriose é possível concluir que o último é o ligando que maior grau de saturação está a receber por parte do receptor molecular, sendo este último putativamente o ligando ótimo de reconhecimento por parte do local de ligação na AraN.

II.3.3.4.4. Interação com arabinotetraose

Na figura II.21 é apresentado o espectro de STD-RMN obtido para a arabinotetraose na presença de AraN.

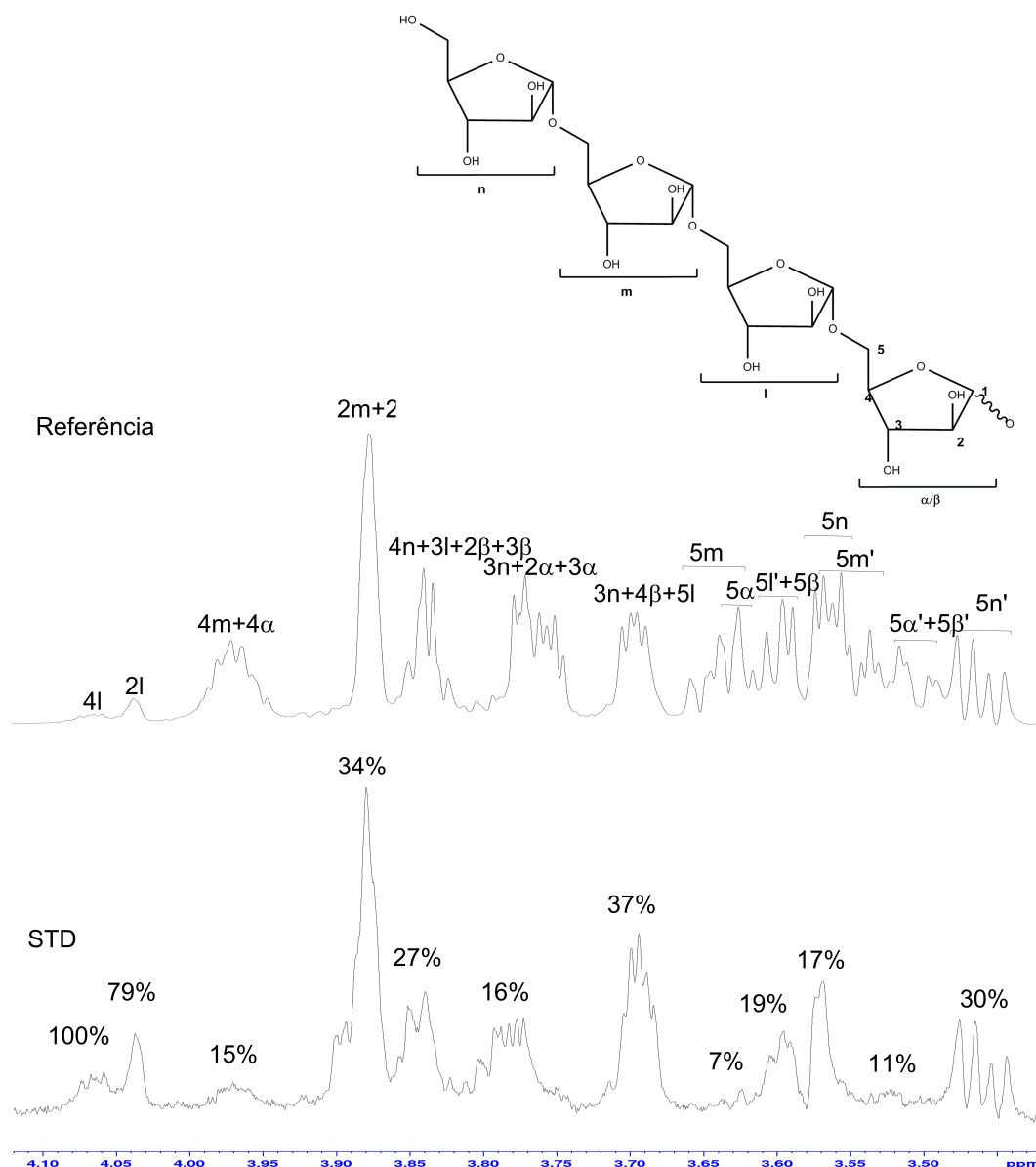


Figura II.21. Espectro de STD-RMN da arabinotetraose com AraN. Na parte superior da figura encontra-se o espectro de referência ^1H -STD off-resonance (I_0), da arabinotetraose, com irradiação a 20000Hz. Em baixo, encontra-se o espectro ^1H -STD (I_{STD}), da arabinotetraose, resultante da diferença entre o espectro I_0 e I_{SAT} , tendo este último sido obtido pela irradiação na região espectral a 150Hz. A aquisição do espectro foi efectuada a 10°C com 2,5 mM de arabinotetraose e 25 μM de AraN.

O objectivo do estudo com este ligando residiu numa tentativa de definir melhor a importância do tamanho do oligossacárido para a promoção da interação proteína-ligando. Relativamente à arabinotriose a arabinotetraose é constituída por mais uma unidade de arabinose tendo-se utilizado o seguinte código para as quatro subunidades; *a* (α/β), *l*, *m* e *n* (sendo a ordem das subunidades da extremidade redutora para a extremidade não redutora).

Como vimos anteriormente na série arabinose, arabinobiose e arabinotriose o parâmetro A_{STD} atinge o máximo para a arabinotriose, pretendeu-se verificar se esta tendência se mantinha com mais uma unidade de arabinose na cadeia. Verifica-se no entanto que para a arabinotetraose são obtidos níveis de A_{STD} equiparáveis aos obtidos para a arabinobiose.

Tal como é possível analisar na figura II.21, para a arabinotetraose a subunidade correspondente ao monómero *l* é aquela que recebe um maior grau de saturação, nomeadamente o protão H4l que apresenta 100% de saturação e o protão H2l que apresenta uma saturação de 79%. Associado a este factor está a depleção do grau de saturação associado ao protão H2n que anteriormente na arabinotriose era de 100% e agora é de 34%. Por outro lado quando comparado o espectro de STD da arabinotriose e arabinotetraose, todas as ressonâncias espectrais associadas à subunidade terminal (*n*) no caso do açúcar de quatro subunidades sofrem uma depleção. Um denominador comum aos estudos efectuados é o de que a subunidade que exibe maior grau de saturação é aquela que se encontra na posição posterior à do monómero na extremidade redutora. Analisados em conjunto estes resultados sugerem que na arabinotetraose a subunidade não redutora (*n*) se deverá encontrar agora mais exposta ao solvente e fora do local de ligação o que poderá explicar a diminuição do A_{STD} .

II.3.3.4.5. Interação com arabinopentaose

Na figura II.22 é apresentado o espectro de STD-RMN obtido para a arabinopentaose na presença de AraN.

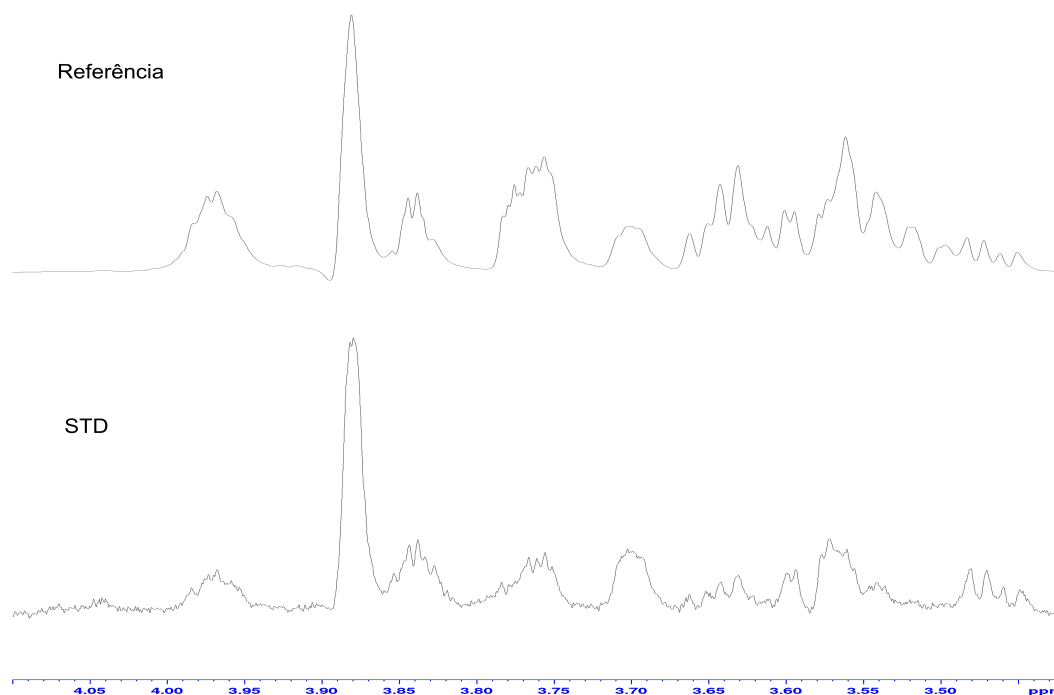


Figura II.22. Espectro de STD-RMN da arabinopentaose com AraN. Na parte superior da figura encontra-se o espectro de referência ^1H -STD *off-resonance* (I_0), da arabinopentaose, com irradiação a 20000Hz. Em baixo, encontra-se o espectro ^1H -STD (I_{STD}), da arabinopentaose, resultante da diferença entre o espectro I_0 e I_{SAT} , tendo este último sido obtido pela irradiação na região espectral a 150Hz. A aquisição do espectro foi efectuada a 10°C com 2,5 mM de arabinopentaose e 25 μM de AraN.

Através da análise da figura II.22, é possível observar presença de sinais no espectro I_{STD} , que é conclusivo da interação com a AraN.

Assim podemos concluir que a AraN apresenta um comportamento de reconhecimento de arabinoligossacáridos de 2 até pelo menos 5 subunidades, apresentando a proteína reconhecimento ótimo da arabinotriose. Ao contrario do observado em CBM's nas quais os carboidratos interagem maioritariamente com a região central do açúcar ficando as extremidades redutoras e não redutoras fora do local de ligação [124, 125], aqui há uma clara preferência pela região redutora e central do açúcar ficando a não redutora fora do local de ligação

Os resultados obtidos até aqui podem ser utilizados para postular um modo recorrente de ligação de arabinoligossacáridos na AraN. Com base nos epítomos e A_{STD} poderemos assumir que existem três locais de interação na cavidade de ligação da proteína denominados de S1, S2 e S3 respectivamente. O reconhecimento parece depender da existência de uma unidade de arabinose a seguir à unidade redutora, o que justifica o facto de a arabinose não ser reconhecida e a arabinobiose já o ser. O fato da unidade a seguir à unidade redutora ser aquela que recebe mais saturação também parece suportar a importância desta segunda unidade para o reconhecimento. Por outro lado uma eficiente interação da unidade não redutora (n) parece ser determinante para a afinidade, visto que há uma reconhecimento ótimo da arabinotriose face à arabinobiose. Assim, à medida que o arabinoligossacárido aumenta de 2 para 3 subunidades há uma ocupação completa dos 3 locais de ligação promovendo-se a interação ótima com a unidade posterior à redutora e com a unidade não redutora. O aumento para 4 subunidades leva no entanto a uma depleção da saturação na extremidade não-redutora do local de ligação e a uma ocupação menos eficiente de um dos locais de ligação por uma unidade não terminal. Este fato é confirmado pelos resultados que mostram que

apenas as três subunidade iniciais a contar da subunidade redutora é que recebem um maior grau de saturação, ou seja pela diminuição de saturação dos prótons H2n e H5n/H5n' com o aumento do tamanho do arabinoligossacárido de 3 para 4 subunidades, evidenciado na figura II.21.

II.3.3.5. Estudo das interações ligando-AraN com celoligossacáridos

Para além, da interação com arabinoligossacáridos foram também testados outros substratos para avaliar o nível de seletividade da AraN no reconhecimento de substratos.

II.3.3.5.1. Interação com celobiose e celotetraose

Ao contrário dos arabinoligossacáridos que são constituídos por monómeros constituídos por cinco carbonos (C₅) estruturalmente arranjados em furano (arabinose) e ligados entre si por ligações glicosídicas α -1,5, os celoligossacáridos são diferentes na medida que são hexaoses (C₆), na forma de pirano compostos por unidades monoméricas de β -D-Glucose, cujas ligações glicosídicas mantidas entre os diferente monómeros são β -1,4 (figura II.23).

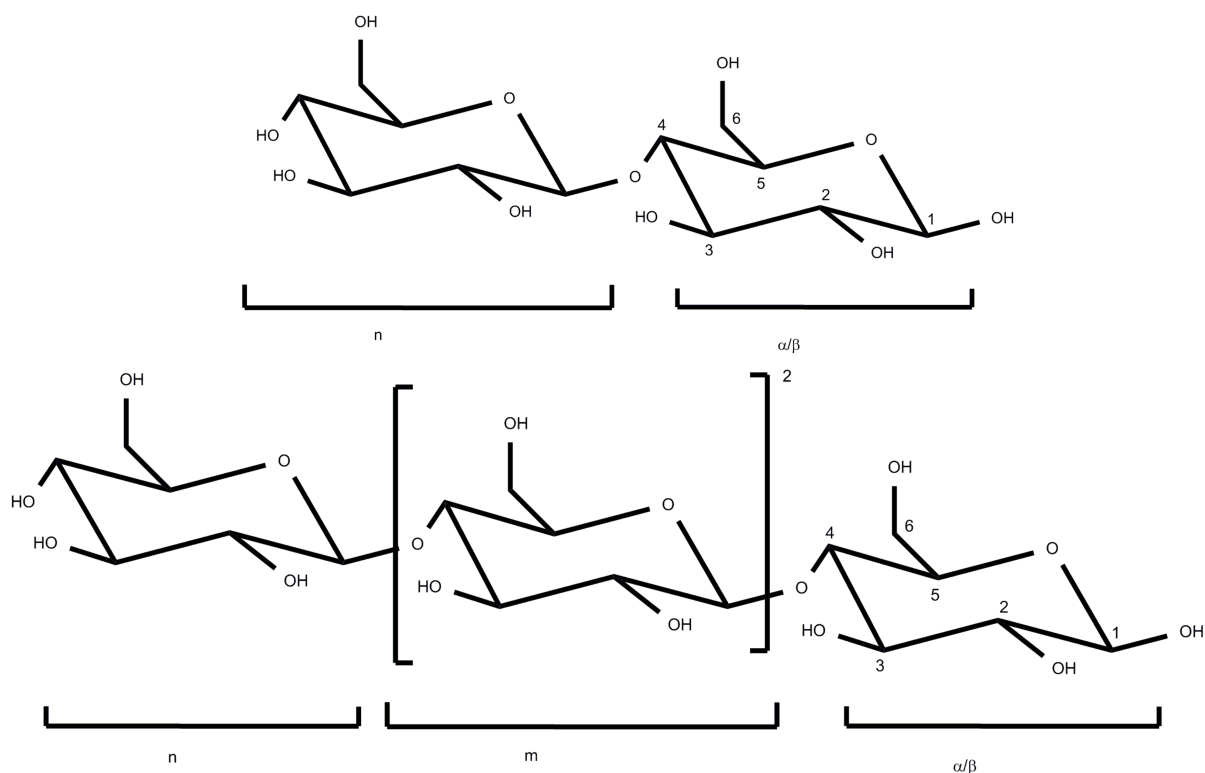


Figura II.23. Estrutura da celobiose e celotetraose. Na parte superior da figura está representada a celobiose, e na região inferior da figura a celotetraose.

Na figura II.24 é apresentado o resultado obtido para a experiência de STD-RMN com a celobiose na presença de AraN.

Como se pode depreender pela presença de sinais no espectro de STD-RMN a AraN revela ter uma maior promiscuidade no reconhecimento de substratos para além do que se conhecia previamente.

Para a celobiose, a região preferencial de reconhecimento que apresenta maior grau de saturação encontra-se no H1n da subunidade não redutora. Pela análise dos resultados é possível

inferir que os prótons da subunidade não redutora que maior saturação recebem comparativamente ao H1n (100%) são o H2n (44%), bem como H6n (41%) e H6n' (33%) dos metilenos.

No que diz respeito à subunidade redutora, estes prótons também se encontram a receber saturação da AraN, sugerindo que esta unidade também possa ser essencial no processo de reconhecimento. Das interações da subunidade redutora as ressonância espectrais que maior grau de saturação apresentam são referentes aos prótons H6β e H5α (57%), H4β (41%), H2α (32%) e H6α (30%).

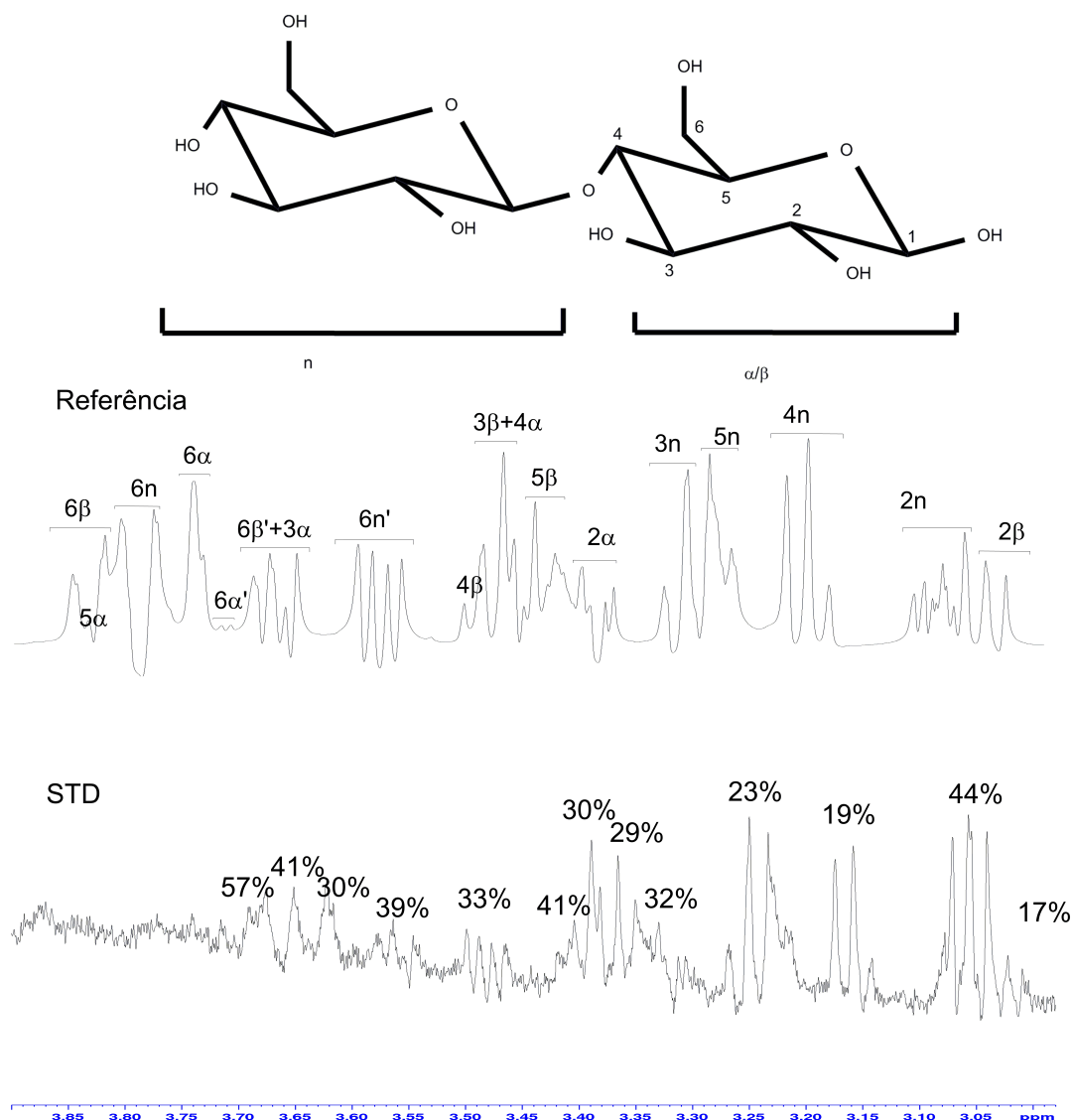


Figura II.24. Espectro de STD-RMN da celobiose com AraN. Na parte superior da figura encontra-se o espectro de referência ^1H -STD *off-resonance* (I_0), da celobiose, com irradiação a 20000Hz. Em baixo, encontra-se o espectro ^1H -STD (I_{STD}), da celobiose, resultante da diferença entre o espectro I_0 e I_{SAT} , tendo este último sido obtido pela irradiação na região espectral a 150Hz. A aquisição do espectro foi efectuada a 10°C com 2,5 mM de celobiose e 25 μM de AraN.

Tal como para o estudo de interação AraN-arabinoligossacáridos, foi utilizado um celoligossacárido com mais de três subunidades, com o objectivo de perceber se há preferência no reconhecimento pelas subunidades centrais ou se este mesmo reconhecimento ocorre nas

extremidades (redutora ou não redutora). Para tal utilizou-se a celotetraose, sendo que o espectro de STD obtido é apresentado na figura II.25.

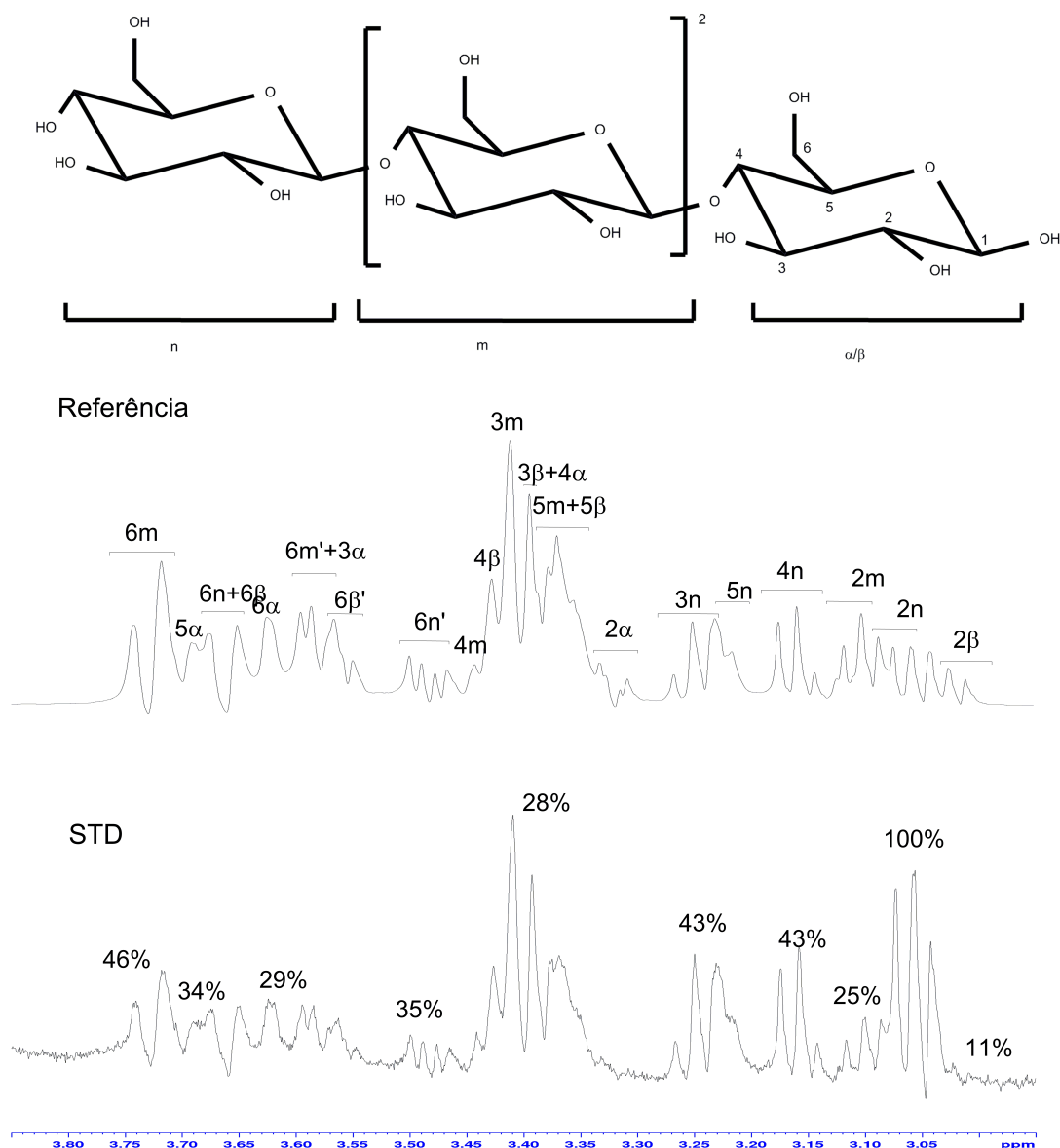


Figura II.25. Espectro de STD-RMN da celotetraose com AraN. Na parte superior da figura encontra-se o espectro de referência ^1H -STD off-resonance (I_0), da celotetraose, com irradiação a 20000Hz. Em baixo, encontra-se o espectro ^1H -STD (I_{STD}), da celotetraose, resultante da diferença entre o espectro I_0 e I_{SAT} , tendo este último sido obtido pela irradiação na região espectral a 150Hz. A aquisição do espectro foi efectuada a 10°C com 2,5 mM de celotetraose e 25 μM de AraN.

A análise dos resultados do estudo de interação com a celotetraose permitem constatar que o protão H2n (100%) da extremidade não redutora apresenta o maior grau de saturação, sendo que os protões H3n, H5n e H4n têm graus de saturação de 43%, apenas superado pelos protões do grupo metilo H6m da subunidade central com 46%.

O facto de existir efeito STD em alguns protões da subunidade redutora, nomeadamente H5 α 34%, bem como o H6 α 29%, pode ser indicativo da necessidade destes em estabilizar o complexo proteína-carboidrato no momento do reconhecimento molecular.

Através da comparação do espectro de STD da celobiose e da celotetraose é possível concluir também que há um aumento no grau de saturação dos prótons associados à subunidade não redutora, nomeadamente o H2n, H3n, H4n e H5n. Assim, e ao contrário do que foi previamente observado para os arabinoligossacáridos, os resultados apontam no sentido dos epítomos de reconhecimento da subunidade não redutora desempenharem um papel importante no reconhecimento de celoligossacáridos por parte da AraN.

II.3.3.6. Screening de outros carboidratos

Com o objectivo de avaliar a interação da AraN com outros oligossacáridos com monómeros na topologia de pirano foram também estudados a manotriose, maltotetraose e a xilopentose, como se encontram representadas na figura II.26.

A 1,4- α -maltotetraose é composta por seis unidades de α -D-glucose (hexose) na conformação de pirano, estando estas subunidades ligadas entre si por ligações glicosídicas α -1,4 (contrariamente aos celoligossacáridos). Por seu turno a 1,4- β -xilopentose é um polímero constituído por cinco subunidades da unidade monomérica de β -D-xilose (pentose), na conformação de pirano. As unidades estão unidas por ligações glicosídicas β -1,4. Em todos estes carboidratos, tal como nos celoligossacáridos os hidroxilos substituintes encontram-se nas suas posições equatoriais. Finalmente a 1,4- β -manotriose é uma hexose composta por três subunidades de β -D-Manose na forma de pirano, polimerizadas por ligações glicosídicas β -1,4.

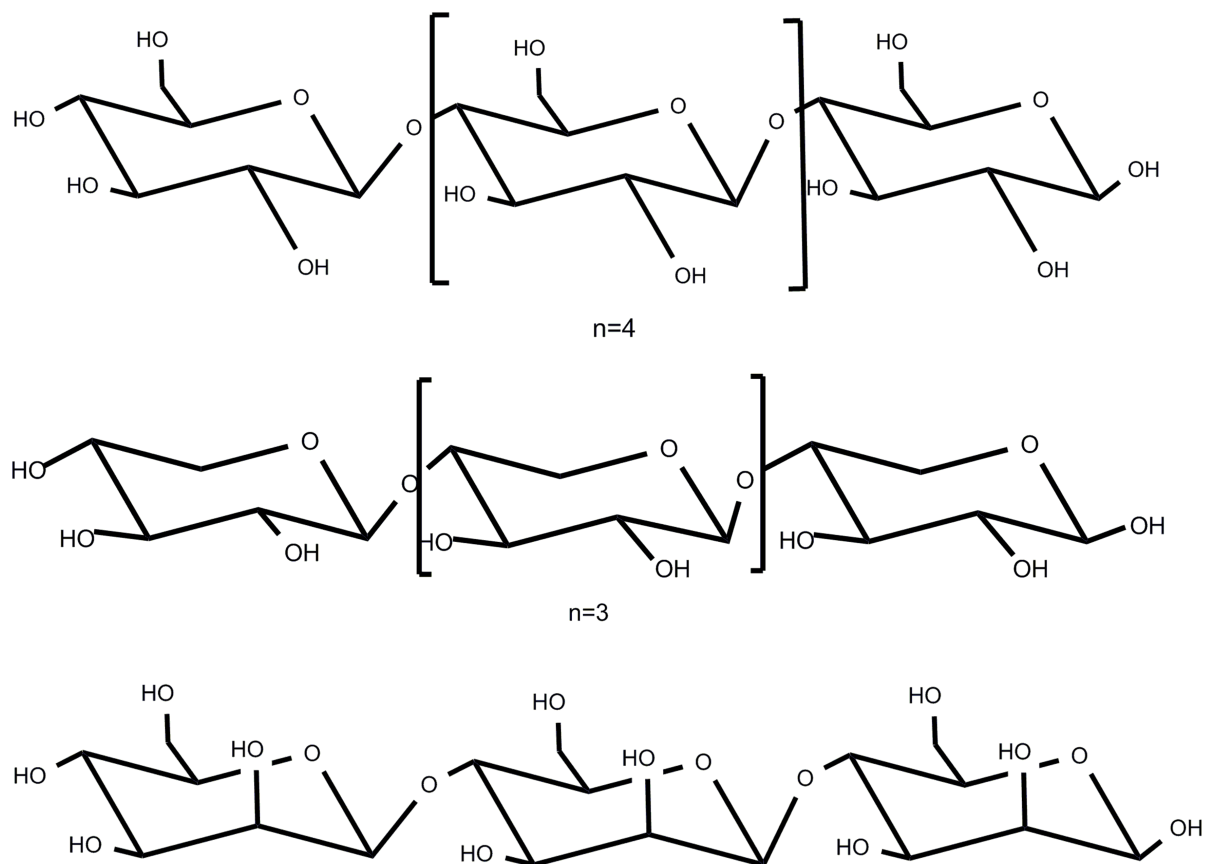


Figura II.26. Estrutura da maltotetraose, xilopentose e manotriose respectivamente.

Nas figuras II.27, II.28 e II.29 encontram-se os resultados dos estudos de STD para a maltohexaose, xilopentaose e manotriose respectivamente.

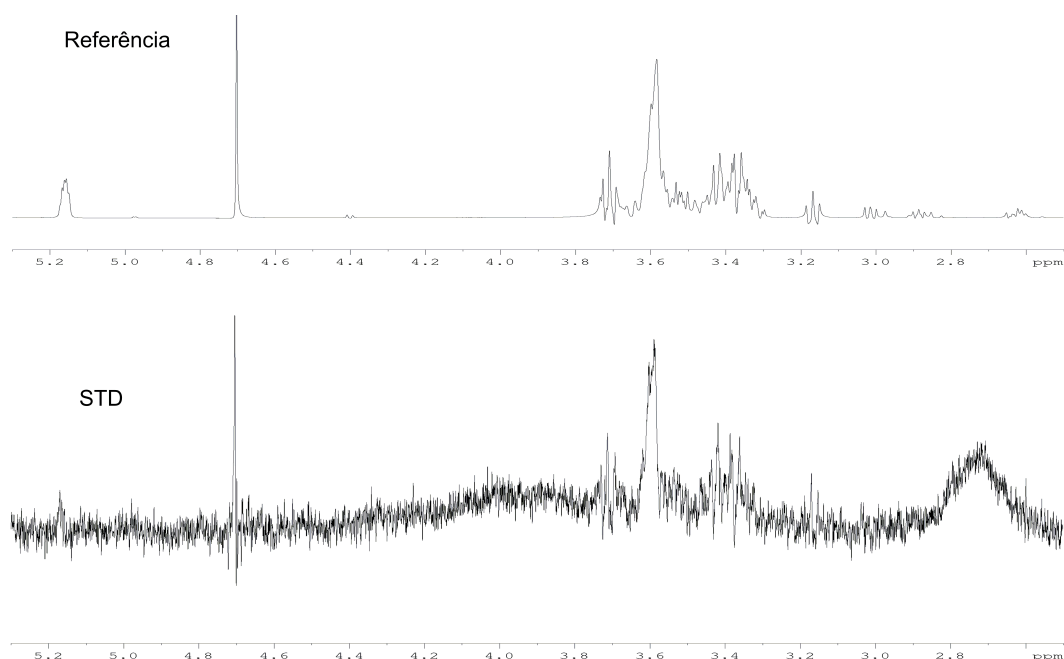


Figura II.27. Espectro de STD-RMN da maltohexaose com AraN. Na parte superior da figura encontra-se o espectro de referência ^1H -STD *off-resonance* (I_0), da maltohexaose, com irradiação a 20000Hz. Em baixo, encontra-se o espectro ^1H -STD (I_{STD}), da maltohexaose, resultante da diferença entre o espectro I_0 e I_{SAT} , tendo este último sido obtido pela irradiação na região espectral a 150Hz. A aquisição do espectro foi efectuada a 10°C com 2,5 mM de maltohexaose e 25 μM de AraN.

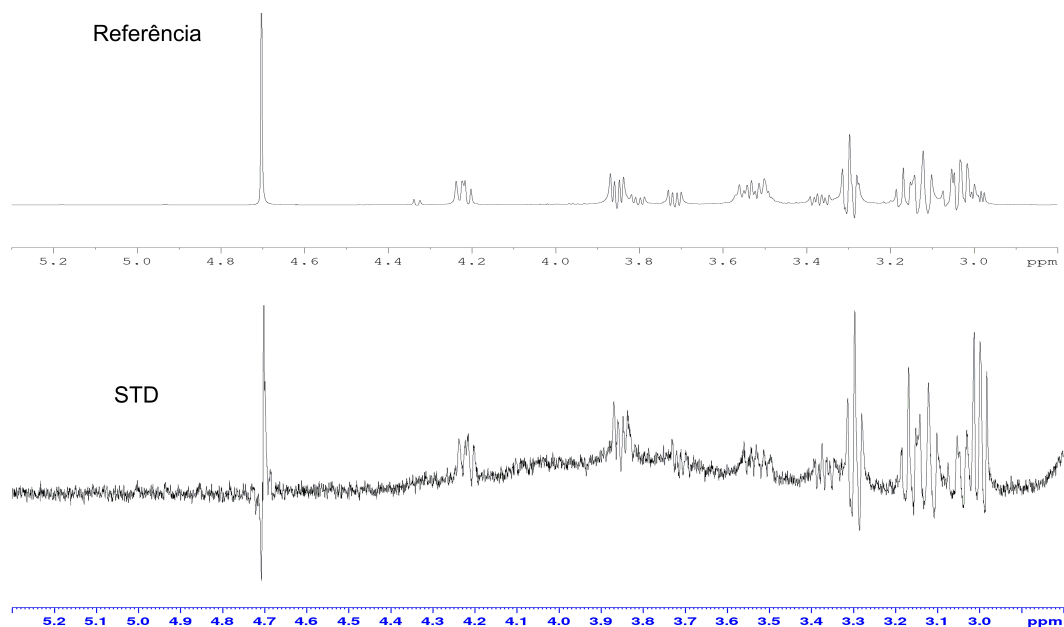


Figura II.28. Espectro de STD-RMN da xilopentaose com AraN. Na parte superior da figura encontra-se o espectro de referência ^1H -STD *off-resonance* (I_0), da xilopentaose, com irradiação a 20000Hz. Em baixo, encontra-se o espectro ^1H -STD (I_{STD}), da xilopentaose, resultante da diferença entre o espectro I_0 e I_{SAT} , tendo este último sido obtido pela irradiação na região espectral a 150Hz. A aquisição do espectro foi efectuada a 10°C com 2,5 mM de xilopentaose e 25 μM de AraN.

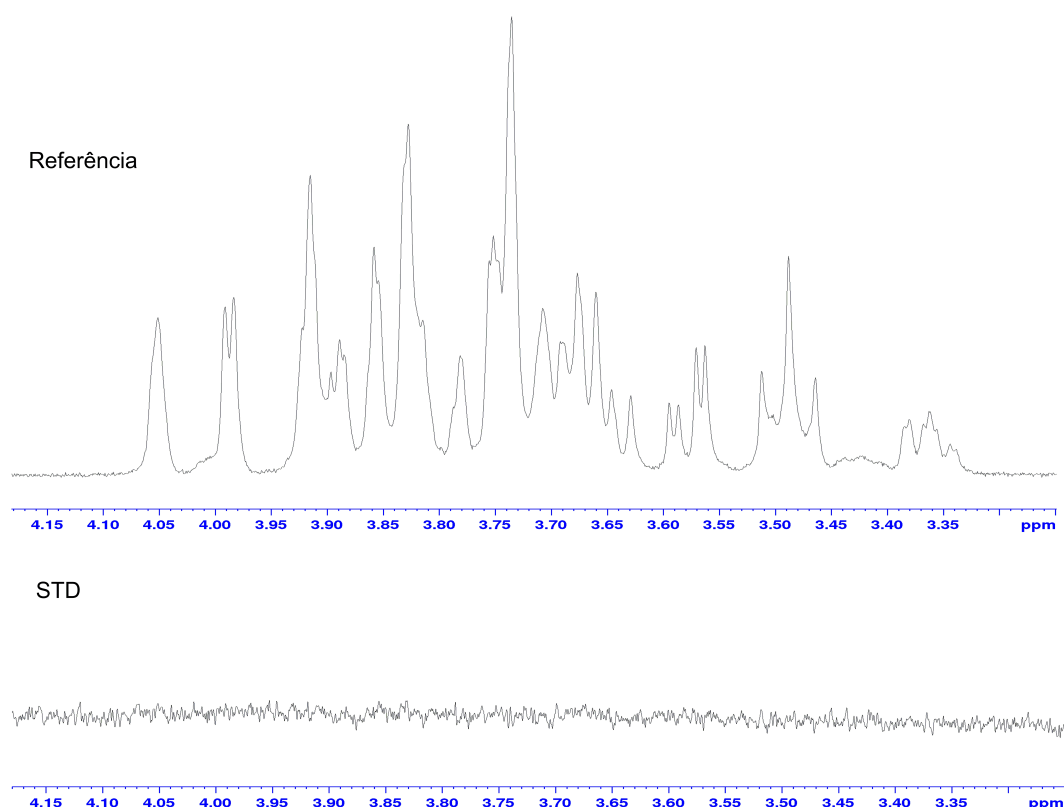


Figura II.29. Espectro de STD-RMN da manotriose com AraN. Na parte superior da figura encontra-se o espectro de referência ^1H -STD *off-resonance* (I_0), da manotriose, com irradiação a 20000Hz. Em baixo, encontra-se o espectro ^1H -STD (I_{STD}), da manotriose, resultante da diferença entre o espectro I_0 e I_{SAT} , tendo este último sido obtido pela irradiação na região espectral a 150Hz. A aquisição do espectro foi efectuada a 10°C com 2,5 mM de manotriose e 25 μM de AraN.

Através da análise das figuras II.27 e II.28 podemos concluir que também estes carboidratos (maltohexaose e xilopentaose) recebem saturação da proteína, no entanto não tendo sido feita a atribuição das ressonâncias espectrais não é possível proceder-se ao mapeamento de epítomos, sendo também essencial perceber se estes STD's são resultantes de interações não específicas, ou seja interações que não sejam resultantes do processo de reconhecimento do ligando pelo local de ligação. Para responder a esta última questão podem ser feitos ensaios de STD-RMN destes carboidratos em condições de competição com um ligando específico, como a arabinotriose. Estes resultados são apresentados na próxima secção.

II.3.3.7. Interação em condições de competição: arabinotriose/maltohexaose e arabinotriose/celotetraose

Na figura II.30 encontra-se o estudo de competição de STD-RMN efectuado com a AraN na presença de arabinotriose e maltohexaose na proporção estequiométrica de 1:1.

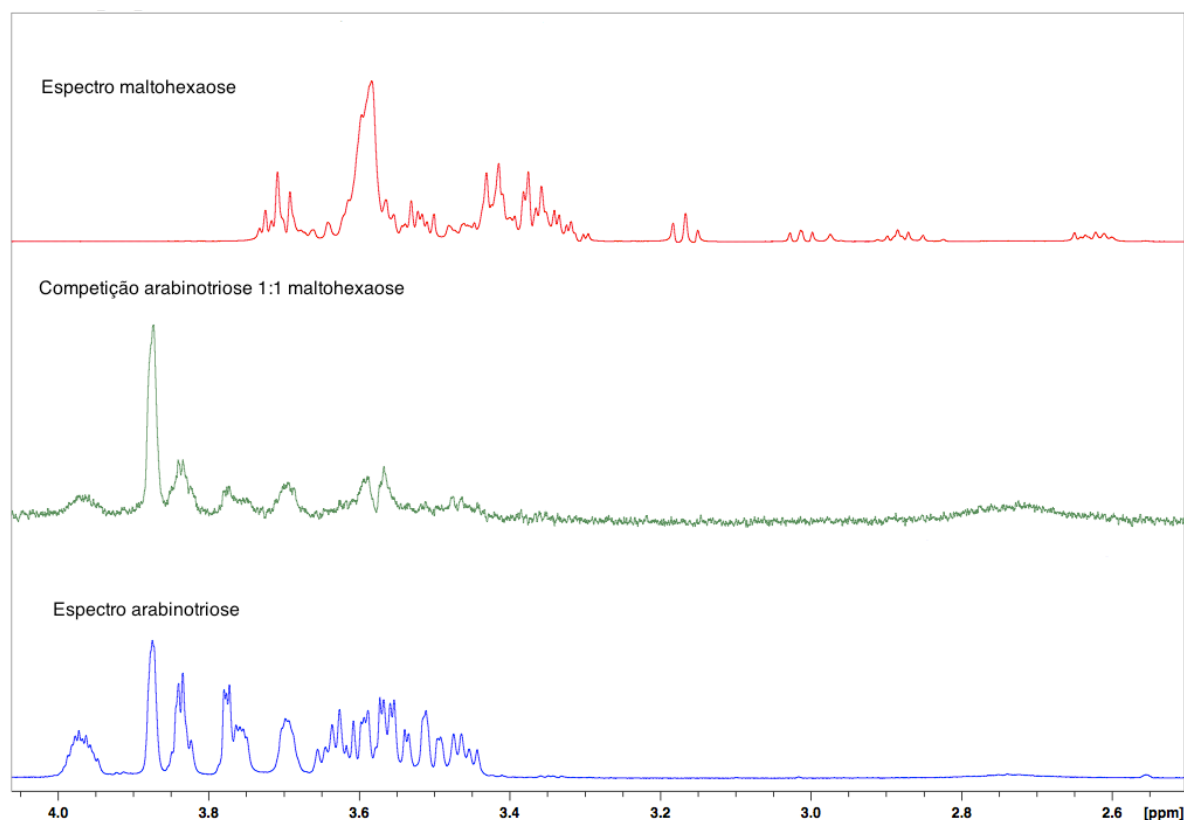


Figura II.30. Espectro de STD-RMN em condições de competição da arabinotriose/manotriose com AraN. Na parte superior da figura encontra-se o espectro de referência ^1H -STD da maltohexaose (traço vermelho). No centro, encontra-se o espectro ^1H -STD (I_{STD}), da competição arabinotriose 1:1 maltohexaose (traço verde), resultante da diferença entre o espectro I_0 e I_{SAT} , tendo este último sido obtido pela irradiação na região espectral a 150Hz, com o espectro I_0 a ser obtido pela irradiação do espectro na zona de 20000 Hz. Na região inferior está representado o espectro ^1H -STD da arabinotriose (traço azul). A aquisição do espectro foi efectuada a 10°C, sendo que a concentração de AraN foi de 25 μM , e os carboidratos de 1,25 mM na competição e 2,5 nos espectros sem competição.

Pela análise do espectro de STD é possível observar que este apenas apresenta ressonâncias espectrais provenientes de prótons da arabinotriose. Assim, pode-se concluir que por um lado o STD evidenciado na figura II.27, no qual são possíveis observar alguns sinais da maltohexaose, não deverão ser provenientes de interações não específicas e que a afinidade da arabinotriose é muito superior que a afinidade da maltohexaose ocupando o mesmo local de ligação na AraN e não permitindo a ligação da maltohexaose.

Na figura II.31, em analogia com o estudo anterior foi adquirido um espectro de STD desta feita com a AraN na presença de arabinotriose e celotetraose na proporção estequiométrica de 1:1.

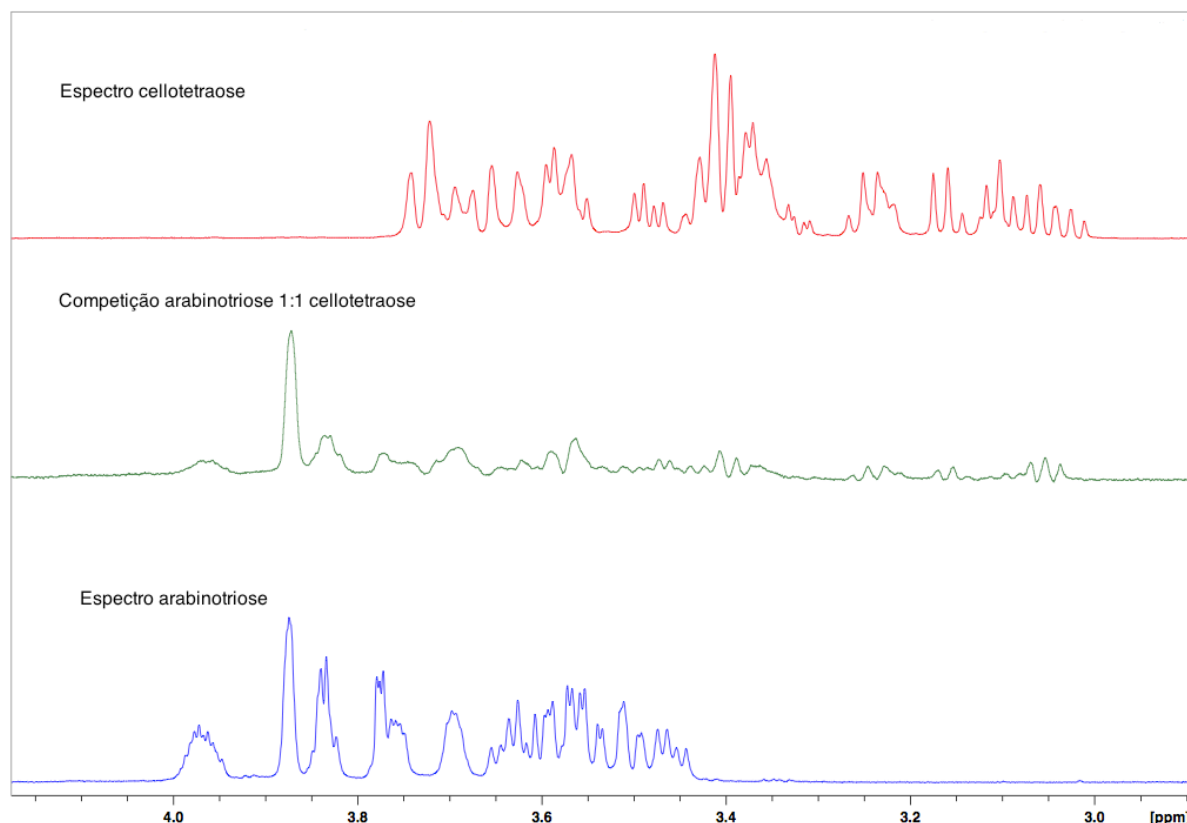


Figura II.31. Espectro de STD-RMN em condições de competição da arabinotriose/celotetraose com AraN. Na parte superior da figura encontra-se o espectro de referência ^1H -STD da celotetraose (traço vermelho). No centro, encontra-se o espectro ^1H -STD (I_{STD}), da competição arabinotriose 1:1 celotetraose (traço verde), resultante da diferença entre o espectro I_0 e I_{SAT} , tendo este último sido obtido pela irradiação na região espectral a 150Hz, com o espectro I_0 a ser obtido pela irradiação do espectro na zona de 20000 Hz. Na região inferior está representado o espectro ^1H -STD da arabinotriose (traço azul). A aquisição do espectro foi efectuada a 10°C, sendo que a concentração de AraN foi de 25 μM , e os carboidratos de 1,25 mM na competição e 2,5 nos espectros sem competição.

Neste caso é possível observar um claro aparecimento de ressonâncias espectrais de ambos os carboidratos. Assim pode-se formular a hipótese que para além do reconhecimento de arabinoligossacáridos por parte de AraN, esta também apresenta capacidade de reconhecimento de celoligossacáridos.

A análise do espectro de STD sugere também que a afinidade da celotetraose para a AraN será da mesma ordem de grandeza da arabinotriose para o mesmo receptor. Esta análise qualitativa carece confirmação quantitativa através dos A_{STD} dos dois carboidratos. Uma vez que os carboidratos estão presentes em condições estequiométricas, o que apresentar sinais com maior A_{STD} será aquele cujos prótons estão em maior proximidade do receptor. No entanto um procedimento mais correto para avaliar a diferença de afinidades é a realização de experiências de STD-RMN para diferentes estequiometrias dos dois carboidratos. A variação da resposta do STD de um carboidrato à medida que se varia a concentração do outro é uma indicação de que os carboidratos competem para o mesmo local de ligação e se um dos K_d for conhecido o nível dessa variação pode permitir uma estimativa do K_d do segundo. Esta foi a abordagem seguida tendo sido realizadas duas experiências, uma com excesso de celotetraose e outra com excesso de arabinotriose.

Na figura II.32, estão representados os estudos de STD para a competição da arabinotriose e celotetraose em diferentes proporções estequiométricas.

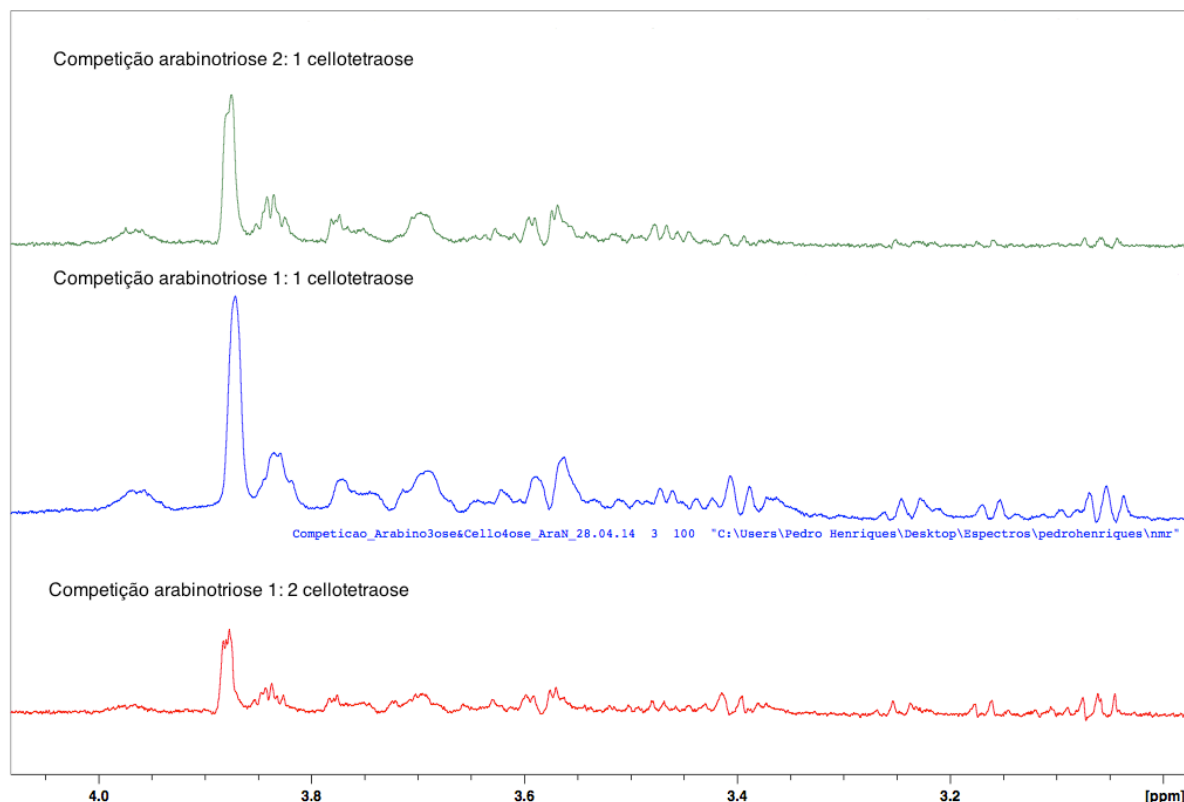


Figura II.32. Espectro de STD-RMN em condições de competição da arabinotriose/manotriose com AraN. Na parte inferior da figura encontra-se o espectro ^1H -STD da celotetraose (traço vermelho), da competição 1:2. No centro, encontra-se o espectro ^1H -STD (I_{STD}), da competição arabinotriose 1:1 celotetraose (traço azul), resultante da diferença entre o espectro I_0 e I_{SAT} , tendo este último sido obtido pela irradiação na região espectral a 150Hz, com o espectro I_0 a ser obtido pela irradiação do espectro na zona de 20000 Hz. Na região superior está representado o espectro ^1H -STD da arabinotriose (traço verde), da competição 2:1. A aquisição do espectro foi efectuada a 10°C, sendo que a concentração de AraN foi de 25 μM , e os carboidratos de 1,7/0,8 mM (2:1) ou 1,3 mM (1:1) na competição e 2,5 nos espectros sem competição.

Através da análise da figura II.32, pode-se concluir que a arabinotriose e a celotetraose competem pelo mesmo local de ligação na AraN e que o processo de reconhecimento molecular apresenta a mesma ordem de grandeza (na secção VI.4. *Espectros suplementares* encontram-se os espectros de STD-NMR com a sobreposição do espectro 1:1 com cada um dos excessos estequeométricos).

Assim podemos afirmar que a interação AraN-celoligossacárido registada na figura II.24 e II.25, não será em principio resultante de interações inespecíficas, mas sim de um reconhecimento por parte da AraN.

II.3.4. Construção de modelos 3D para análise e interpretação de interações

Quando a estrutura do receptor molecular é conhecida a análise dos epítomos obtidos pelas experiências de STD-RMN pode ser complementada através de modelos tridimensionais dos complexos obtidos por metodologias computacionais de *docking* molecular. Os modelos assim obtidos permitem outro nível de entendimento e racionalização do processo de reconhecimento molecular, uma vez que se obtém também informação das interações do ponto de vista do receptor, apesar destas serem obtidas através de simulações *in silico*.

No caso da AraN a sua estrutura não é conhecida, no entanto, é sabido que as SBP's apesar de não apresentarem uma homologia de sequência primária de aminoácidos, apresentam todas elas entre si um grande homologia ao nível a estrutura terciária [53]. Com base neste fato decidiu-se recorrer a ferramentas informáticas para determinar um modelo de homologia para a estrutura da AraN. Esse modelo foi posteriormente utilizado para estudos de *docking* molecular.

II.3.4.1. Modelo de homologia estrutural 3D da AraN

O modelo obtido pelo I-TASSER [126] foi obtido de acordo com o procedimento descrito nos materiais e métodos (secção II.2.7) que obedecem essencialmente a três etapas: *Threading*, *ASSEMBLY* e *Refinement*.

Na primeira etapa, regiões da sequência primária da proteína são alinhadas com uma biblioteca de estruturas depositadas no Protein Data Bank (PDB) (este primeiro alinhamento é feito tendo por base um *cut-off* de 70% de identidade). A principal função desta primeira etapa reside em encontrar possíveis motivos estruturais de sequência secundária, combinando para isso diferentes algoritmos de alinhamento.

Na segunda-etapa os alinhamentos são classificados segundo regiões alinhadas e regiões não alinhadas. As primeiras são excisadas e utilizadas no modelo ao passo que as segundas irão sofrer um processo de modelação por simulação *ab initio*. O processo de montagem da estrutura é efectuado por simulações Monte Carlo, sendo que as trajetórias obtidas na estrutura tri-dimensional são agrupadas. Através destes agrupamentos são gerados centroídes, tendo por base a média de todas as estruturas agrupadas.

Finalmente este processo é repetido iterativamente com o intuito de refinar a estrutura, bem como remover impedimentos estereoquímicos. As pontuações obtidas para o modelo da AraN encontram-se na tabela II.14 e o modelo está representado na figura II.33.

Tabela II.14. Tabela com as pontuações obtidas do modelo de homologia da AraN. Associado ao modelo de homologia estão parâmetros que estimam a qualidade do modelo, tais como o C-Score, TM-Score e RMSD.

C-Score	0,99
TM-Score	0,85±0,08
RMSD	4,8±3,2Å

O parâmetro C-Score (*confidence score*), dá uma estimativa do grau de confiança que o servidor tem relativamente ao modelo obtido. Esta pontuação é dada consoante o grau de alinhamento global adoptado entre a estrutura obtida comparativamente às estruturas homólogas. A função que dá origem ao C-score é constituída por três factores. Os dois primeiros factores estão relacionados com a convergência das estruturas obtidas no agrupamento (*cluster*) durante o processo de *Assembly*. A terceira função está associada com a qualidade dos alinhamentos efectuados aquando do processo de *Threading*. Por seu turno o RMSD, está associado à média das distâncias entre todos os pares de resíduos entre as estruturas homólogas e o modelo. Geralmente um bom C-score pode não estar associado a um bom RMSD. Isto pode ser explicado pelo facto do RMSD ser sensível a regiões desordenadas, ou seja, um grande erro local na estrutura modelo irá dar origem a um RMSD alto, pese embora que a topologia da proteína modelo apresente uma elevada homologia com as proteínas homólogas (C-score alto). A pontuação C-score é avaliada entre -5 e 2, sendo que quanto maior for este valor maior é a validade da estrutura [126]

Finalmente o TM-score é um género de RMSD ponderado. Ou seja, como vimos anteriormente o RMSD é muito sensível a zonas desordenadas locais, assim o TM-score apresenta um índice de pontuação que dá maiores pontuações a regiões ordenadas e penalizações pequenas a regiões locais desordenadas. O TM-score é dado entre 0 e 1, com os melhores modelos a apresentarem pontuações próximas de 1 [126, 127].

Assim e analisando o C-score e o TM-score estes apresentam valores que levam a concluir que o modelo gerado apresenta uma topologia muito semelhante à topologia adoptada por proteínas homólogas, sendo que para isto em muito contribui o facto das proteínas SBP's apresentarem uma elevada conservação da estrutura tridimensional, entre diferentes famílias da mesma proteína para diferentes sistemas biológicos. No entanto, actualmente é insubstituível a obtenção de estruturas tridimensionais ora por RMN ou por cristalografia de raio-X, pois apenas estas apresentam uma resolução para estudos de interacção. É espectável que o local de ligação da proteína gerada por homologia não seja 100% fidedigna comparativamente à estrutura real em solução.

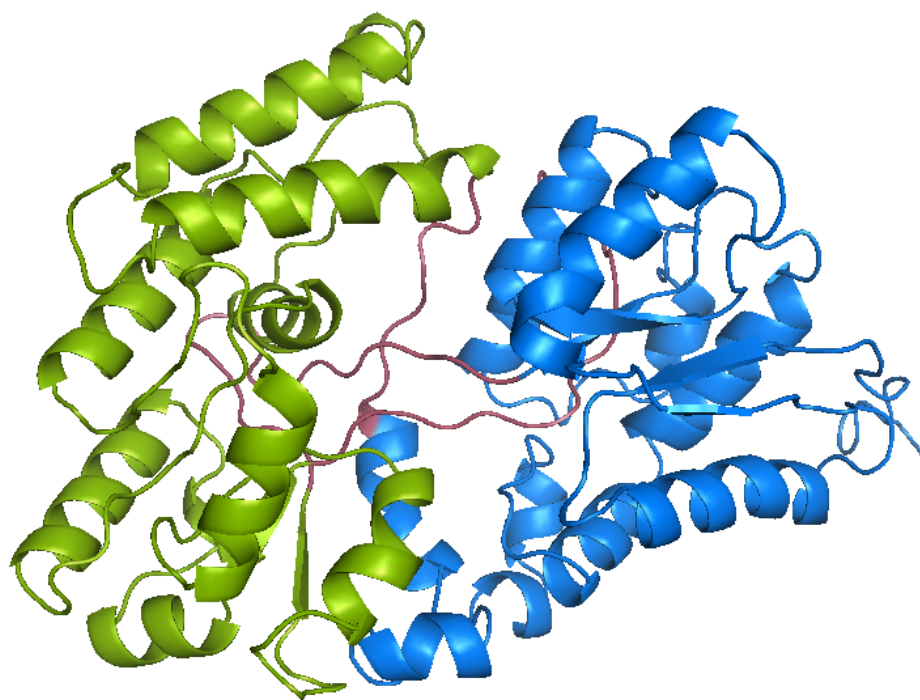


Figura II.33. Representação da AraN obtida no servidor I-TASSER. O domínio C-terminal está representado a verde, o domínio N-terminal encontra-se representado a azul. As três sequencias peptídicas que unem estes dois domínios encontram-se representadas a vermelho.

A AraN obtida pelo modelo I-TASSER é provavelmente uma SBP do tipo I, visto que apresenta dois domínios ambos separados em três sequencias de aminoácidos. Cada sequencia de aminoácidos encontra-se separada por um sequencia peptídica que irão formar as dobras (*hinges*) responsáveis pelo mecanismos *hinge bending* [55]. Em baixo encontra-se a sequencia peptídica colorida com o código da figura II.33:

MCSAEKSSGKSGETELTFWTFNGLHEQFYVEMVKEWNKKYPDRKIKLNTVVYPYQGMHDNLSISLIAGEGVDPDI
 ADVELARFSNFLKGSDIPLADLTPLIEKDRDKFVEARLTL**YSKNGKLYGLDTHVGTTFVMFY**NMDVMKKAGVNPD
 DIKTWDDYHKAGQKVRKVTGKPMGTVETNDSATFLSMISQQNSGYFDKNGKLILNNDTNVKT**LQYLKDMINDKT**
 MIPAPGGGGHHSEYYGFMNQGAASVLMPIWYMGRFIDYMPDLKGK**IAIRPLPAWKEGGDRSAGLGGTATVVP**
 KQSKHVELAKEFLAFAKGSEEGNKKLWSVLGFDPLRWDVWSSKELKEKNKYTDYFQNGTGIFSVLLDIKD**INP**
IYLHEDFAKASDLVNRSVLFDALKSQQKTPKQALDRAAGELKQKHLE

II.3.5 Localização do local de ligação da AraN por alinhamento estrutural

Para a localização do local de ligação do substrato à AraN optou-se por uma estratégia que permite alinhar sequências utilizando informação estrutural de sequências previamente selecionadas. Para este efeito utilizaram-se sequências de proteínas com estrutura depositada em complexo com o seu ligando como discutido mais adiante. Neste caso o servidor utilizado foi o Expresso [105].

Resumidamente o Expresso irá fazer o BLASTing de todas as sequências inseridas no programa contra uma biblioteca de proteínas como é o caso do PDB, com o intuito de encontrar modelos apropriados às sequências. Assim que cada modelo é encontrado para cada sequência primária de cada proteína, é produzido um alinhamento baseado em motivos estruturais conservados,

para alinhar todos os modelos estruturais previamente encontrados através do PDB, criando com isso uma biblioteca (associando motivos estruturais a regiões de sequência primária).

Assim e garantindo que uma grande percentagem das proteínas utilizadas na sequenciação através do Expresso se encontre depositada no PDB, irá obter-se um produto final no qual as sequências se irão encontrar alinhadas segundo motivos estruturais conservados.

Uma escolha criteriosa das estruturas a utilizar neste procedimento é crucial para que os dados obtidos sejam fiáveis para posterior interpretação no contexto do reconhecimento molecular de ligandos por parte da AraN.

Tal como se encontra reportado na literatura existem regiões específicas no local de ligação das SBP responsáveis pelo reconhecimento do substrato. As SBP são constituídas essencialmente por dois domínios o N-terminal e o C-terminal, sendo que o reconhecimento inicial do substrato no caso da MBP ocorre através do domínio C-terminal. Este reconhecimento dá-se através de interações hidrofóbicas, em particular pelo *stacking* com resíduos cujas cadeias laterais apresentam anéis aromáticos. Em grande parte este reconhecimento é efectuado por triptofanos e tirosinas. Após um primeiro reconhecimento do substrato pelo C-terminal (domínio II), o domínio N-terminal (domínio I) irá aproximar-se do C-terminal estabilizando o complexo através da formação de pontes de hidrogénio com o mesmo [107].

No caso da MBP de *Escherichia coli* existem três resíduos proeminentes no reconhecimento de maltotriose sendo estes o Y155, W230 e W340. Estes resíduos foram também encontrados no local de reconhecimento de ciclodextrinas em *Thermoactinomyces vulgaris*, que neste caso são Y175, W250 e W360 [108, 128]. Suzuki *et al* [109] demonstraram através da cristalização de proteínas de ligação a galacto-N-biose (GBP) e lacto-N-biose (LBP), que estas duas proteínas reconhecem estes carboidratos com um mecanismo semelhante entre elas, sendo que existem dois triptofanos conservados responsáveis pelo reconhecimento inicial (W257 e W278), bem como cinco aminoácidos polares conservados no processo de reconhecimento (R49, Q105, D154, E203, S316) [109]. No entanto, a GBP e LBP reconhecem os substratos através de mecanismos mediados por águas que irão fazer a ponte entre os resíduos polares e os hidroxilos do carboidrato, ao passo que no caso da MBP esta apresenta uma maior hidrofobicidade no seu local de ligação sendo o complexo estabilizado em grande parte por interações hidrofóbicas, revelando assim dois mecanismos dispares no que toca ao reconhecimento molecular em SBP's.

Como foi previamente discutido os resultados das experiências de STD-RMN, apontam no sentido de haver regiões responsáveis pelo reconhecimento individual de cada subunidade que compõe os oligossacáridos [129]. Cuneo *et al* [110] explora esses locais de ligação para quatro proteínas diferentes: a MBP de *Escherichi coli* [107], a MBP de *Termococcus litoralis* [112], a GBP de *Thermus thermophilus* [114], e a MtBP de *Thermus thermophilus* [110]. Nesse estudo pôde observar-se que existem motivos estruturais (2 loops e 3 hélices) que apresentam resíduos responsáveis pelo reconhecimento de subunidades dos carboidratos em questão. O mecanismo de reconhecimento de maltodextrinas parece ser conservado entre diferentes organismos. Por exemplo, em *Streptococcus penumoniae*, a SBP consegue reconhecer maltoligossacáridos até oito subunidades no entanto o núcleo de reconhecimento apenas consegue albergar três subunidades, sendo este o

açúcar que maior afinidade apresenta para a proteína. Novamente o reconhecimento é efectuado por aminoácidos com cadeia lateral aromática tais como o Y197 (local 2), W273 (local 1) e W384 (local 3) [54]. No local 1 registam-se ainda formação de pontes de hidrogénio com o açúcar mediados por K306 e Y51. A baixa afinidade para a maltose e afinidade máxima para a maltotriose, revela que o local de ligação 3 para com a subunidade não redutora desempenha um papel fulcral no reconhecimento de carboidratos, local este onde se encontra o núcleo das interações van der Waals bem como interações hidrofóbicas CH-*stacking*.

Assim e com o intuito de identificar resíduos putativos de reconhecer a arabinotriose no caso da AraN foram efectuados alinhamentos com as seguintes sequências de proteínas: Maltose BP de *Escherichia coli* (PDB: 3MBP) [107], ciclo/maltodextrinas BP de *Thermoactinomyces vulgaris* (PDB:2ZYO) [108], galacto-N-biose/lacto-N-biose BP de *Bifidobacterium longum* (PDB:2Z8D) [109], maltose BP de *Thermus thermophilus* (PDB:2GH9) [110], GachH BP de *Streptomyces glaycescens* (PDB:3K00) [111], trehalose/maltose BP de *Thermococcus litoralis* (PDB:1EU8) [112], maltodextrin BP de *Pyrococcus furiosus* (PDB:1ELJ) [113], maltoligossacárido BP de *Streptococcus pneumoniae* (PDB:2XD3) [54] e da glucose BP de *Thermus Thermophilus* (PDB:2B3B) [114].

A análise dos alinhamentos de sequência permitiu identificar resíduos putativos que deverão ter uma alta probabilidade de se encontrarem a interagir com a arabinotriose, sendo estes o R81, E175, D178 e W253. Estes resíduos, são responsáveis por exemplo pelo reconhecimento da maltotriose na MBP (R66, W230, E153) [107] e da galacto-N-biose e lacto-N-biose na GL-BP (W278) [109], e de ciclodextrina de *TvuCMBP* (R84, E170 e W250) [108]. Tal como é possível observar nas figura II.34, existem três resíduos conservados ao longo desta família de SBP's, nomeadamente dois triptofanos e uma tirosina, no entanto apenas o triptofano 253 se encontra conservado na AraN.



Figura II.34. Alinhamento das sequências primárias com recurso ao Expresso. As caixas a preto identificam aminoácidos de cadeia lateral aromática envolvidos no reconhecimento de substratos em algumas SBP's.

Ao analisar outros resíduos que se encontram a interagir com o seu substrato foram também encontrados resíduos passíveis de serem responsáveis por efetuar pontes de hidrogénio com a arabinoligossacáridos no caso da AraN, estes resíduos encontram-se identificados na figura II.35. No que toca à MBP de *Escherichia coli* responsável pelo reconhecimento de maltose, maltotriose e maltotetraose, o resíduo R66 encontra-se envolvido na formação de pontes de hidrogénio com subunidades não-redutoras dos substratos acima evidenciados. Nomeadamente no caso da maltose o resíduo R66 forma pontes de hidrogénio com o OH3 e OH4, no caso da maltotriose com o OH2 e OH3 e finalmente na maltotetraose com o OH3. Por outro lado, a proteína *TvuCMBP* de

Thermoactinomyces vulgaris, tem entre outros, o resíduo R84 conservado no processo de reconhecimento de maltotetraose e maltodextrinas.

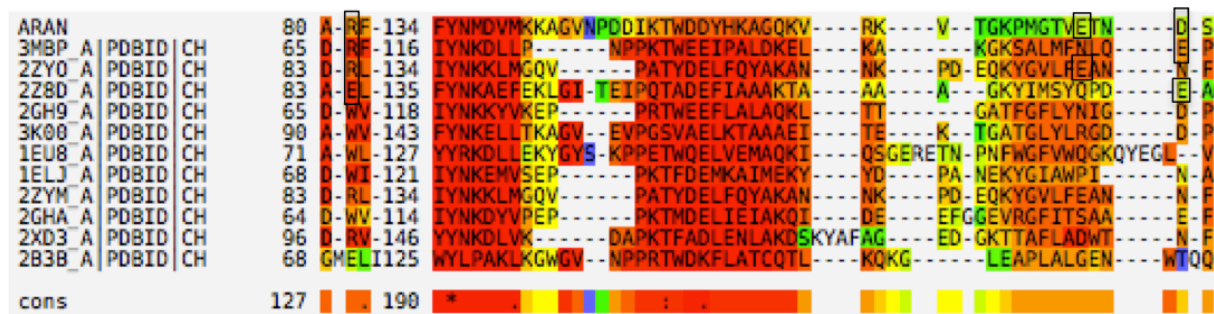


Figura II.35. Alinhamento das sequências primárias com recurso ao Expresso. As caixas identificam aminoácidos envolvidos no reconhecimento do substratos para os SBP's em questão, através da formação de pontes de hidrogénio.

Na figura II.36 é possível observar que alguns aminoácidos da MBP responsáveis pelo reconhecimento da maltotriose encontram-se estruturalmente sobrepostos na AraN. Os aminoácidos D178 (AraN) e E158 (MBP), W253 (AraN) e W230 (MBP) bem como R81 (AraN) e R66 (MBP) encontram-se sobrepostos na mesma região de ambas as proteínas.

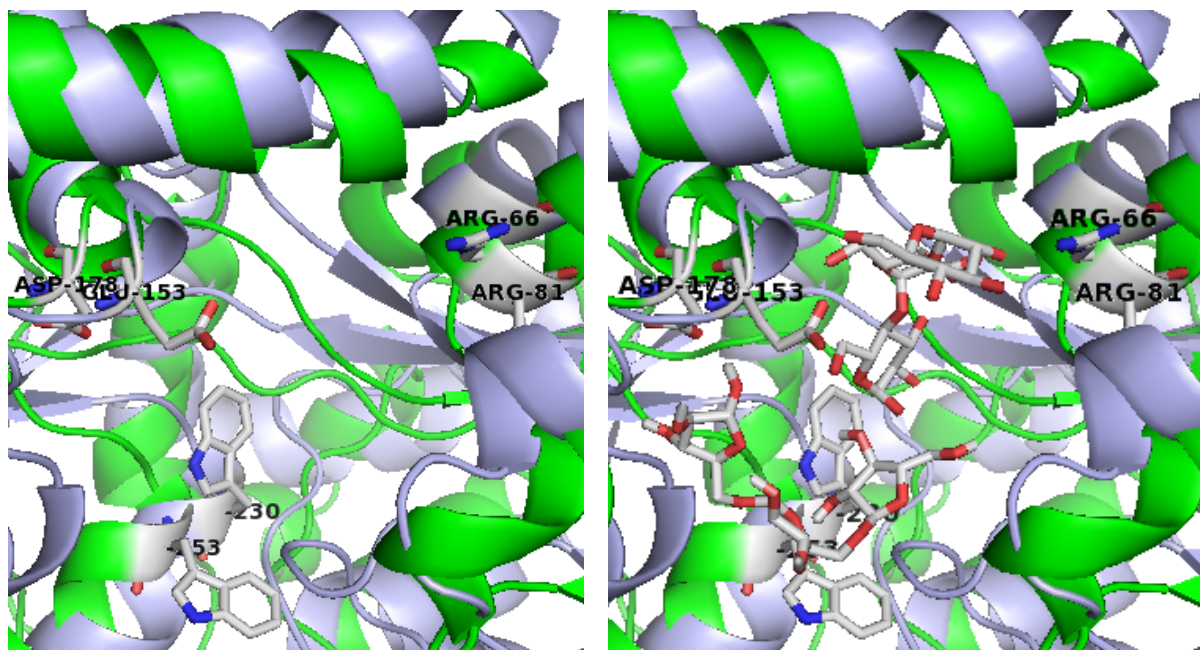


Figura II.36. Alinhamento da estrutura da AraN com a MBP. Neste painel de figuras a AraN (verde) encontra-se alinhada com a MBP (azul). O alinhamento de AraN e MBP foi conseguido com um RMSD global de 3.836Å. Conforme previsto no alinhamento Expresso os aminoácidos D178 (AraN) e E158 (MBP), W253 (AraN) e W230 (MBP), bem como os resíduos R66 (MBP) e R81 (AraN) encontram-se sobrepostos na mesma região de ambas as proteínas, envolvidas no processo de reconhecimento dos respectivos substratos, como é visível na figuras à direita. PDB MBP:3MBP.

Por outro lado o mesmo se observa com a *TvuCMBP*, na qual os aminoácidos D178 (AraN) e N173 (*TvuCMBP*-G4), E175 (AraN) e E170 (*TvuCMBP*-G4), W253 (AraN) e W250 (*TvuCMBP*-G4), bem como os resíduos R81 (AraN) e R84 (*TvuCMBP*) estão estruturalmente conservados em ambas as proteínas.

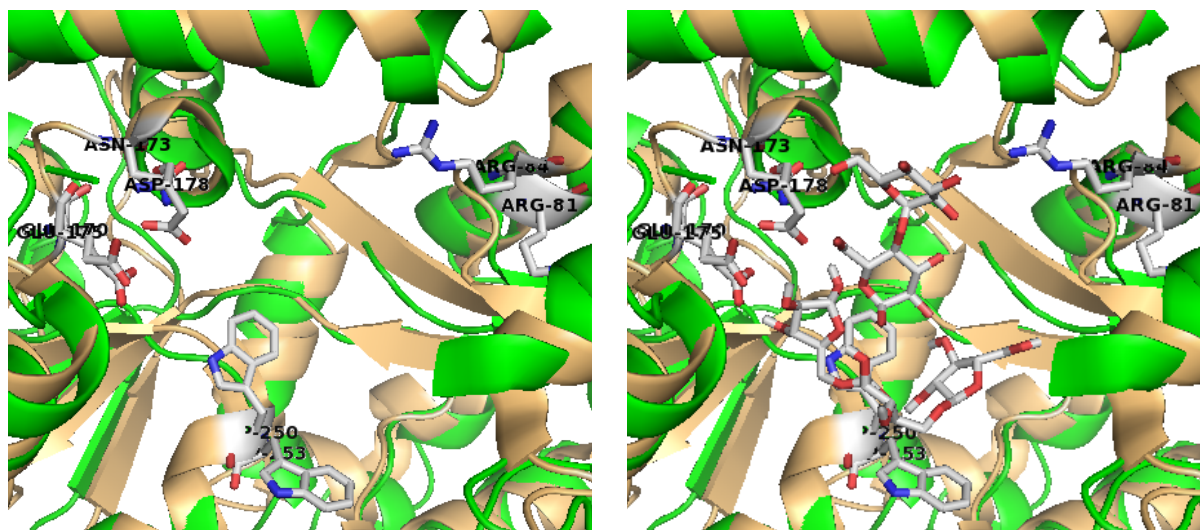


Figura II.37. Alinhamento da estrutura da AraN com a *TvuCMBP-G4*. Neste painel de figuras a AraN (verde) encontra-se alinhada com a *TvuCMBP-G4* (amarelo). O alinhamento da AraN e *TvuCMBP-G4* foi conseguido com um RMSD global de 2.124 Å. Conforme previsto no alinhamento Expresso os aminoácidos D178 (AraN) e N173 (*TvuCMBP-G4*), E175 (AraN) e E170 (*TvuCMBP-G4*), W253 (AraN) e W250 (*TvuCMBP-G4*), bem como os resíduos R84 (*TvuCMBP*) e R81 (AraN) encontram-se sobrepostos na mesma região de ambas as proteínas, envolvidas no processo de reconhecimento dos respectivos substratos, como é visível na figuras à direita. PDB *TvuCMBP-G4*: 2ZY0.

Este estudo permitiu identificar resíduos que podem ser essenciais para o reconhecimento do ligando por parte da proteína AraN. Este mapeamento foi posteriormente utilizado para definir regiões de interação na superfície da proteína para os estudos de *docking* molecular (secção II.3.6.), e para dirigir estudos de mutagenese (secção II.3.7.)

II.3.6. *Docking* molecular

O *docking* molecular é nos dias correntes, com a rápida expansão de áreas como a genómica e proteómica, uma ferramenta vastamente utilizada na previsão de interações biomoleculares, com o objectivo de estudar modos de reconhecimento entre receptores moleculares e outros ligandos (DNA, RNA, proteínas e carbohidratos) [130].

Para os efeitos destes estudos foi utilizado o programa AutoDock, com o objectivo de por um lado estudar a conformação de ligação do ligando no receptor molecular e por outro lado prever quais os resíduos do ponto de vista a AraN putativos de se encontrarem a interagir com a arabinotriose. Para tal o Autodock recorre a um método de procura espacial para encontrar locais de ligação presentes na AraN (*grid-based method*) e utiliza um campo de forças com o objectivo de avaliar as energéticas envolvidas no processo de reconhecimento molecular de cada conformação da arabinotriose.

O processo iterativo de procura das melhores conformações do ligando faz uso de um algoritmo genético Lamarckiano. Numa primeira fase é criada um população de conformações, sendo estas populações sucessivamente mudadas nas gerações seguintes (i.e. alvo de alterações conformacionais), competindo a prole entre si num processo análogo à evolução biológica, ou seja as conformações de mais baixa energia e consequente mais aptas irão ser seleccionadas. A vantagem de ser utilizado um algoritmo Lamarckiano é que em cada processo iterativo, cada

conformação é submetida a um mínimo local energético, sendo esta informação passada à geração seguinte e assim sucessivamente até serem encontradas as melhores conformações [131].

Numa fase final as conformações são organizadas em *clusters* por ordem de energias, sendo que no mesmo *cluster* irão encontrar-se as conformações cujo o RMSD varia em apenas 2.0 Å, caso uma conformação não preencha este último parâmetro ficará num novo *cluster* e assim sucessivamente.

O cálculo da energia livre de ligação é conseguido através da diferença entre a energia livre do ligando e da proteína nas suas formas não-ligadas (equação 4) e da energia livre do complexo proteína-ligando (equação 5). Ou seja, numa primeira fase são avaliadas as energéticas intramoleculares do processo de transição do estado não complexado para o estado complexado para cada uma das moléculas envolvidas sozinhas, e numa segunda fase as energéticas intermoleculares envolvidas no processo de formação do complexo. O primeiro termo da equação 4 é relativo à diferença de energias intramoleculares do estado ligado e desligado do ligando, ao passo que o segundo termo é relativo à diferença de energias intramoleculares do estado ligado e desligado da proteína. O terceiro e último termo está associado à diferença energética intermolecular da formação do complexo [132].

$$\Delta G = (V_{bound}^{L-L} - V_{unbound}^{L-L}) + (V_{bound}^{P-P} - V_{unbound}^{P-P}) + (V_{bound}^{P-L} - V_{unbound}^{P-L} + \Delta S_{conf}) \quad \text{Equação 4}$$

A equação 5 faz uso de cinco factores empíricos envolvidos na formação do complexo proteína-ligando que descrevem as interações van der Waals, pontes de hidrogénio, interações electrostáticas, entropia, e desolvatação.

$\Delta G =$

$$f_{vdW} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + f_{hbond} \times$$

$$\left(\sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + E_{hbond} \right) + f_{elec} \sum_{i,j} \left(\frac{q_i q_j}{\epsilon r_{ij}} \right) + \Delta G_{tor} N_{tor} + f_{sol} \sum_{i,j} (S_i V_j + S_j V_i) e^{\frac{-r_{ij}^2}{2\sigma^2}} \quad \text{Equação 5}$$

II.3.6.1 Clusters obtidos através do AutoDock

A simulação no AutoDock foi executada para gerar dez conformações de arabinotriose no local de ligação da AraN, sendo que estas encontram-se organizadas por energias livres na tabela II.15.

Tabela II.15. Parâmetros retirados da simulação efectuada no AutoDock.

Conformações	1	2	3	4	5	6	7	8	9	10
Energia de ligação	-2,12	-1,79	-2,75	-1,6	-1,4	-1,1	-2,58	-2,07	-0,87	-2,11
Constante de inibição (mM)	28,05	48,69	9,61	66,8	93,73	156,08	12,87	30,58	231,17	28,5
Energia intermolecular	-6,59	-6,27	-7,23	-6,08	-5,88	-5,57	-7,05	-6,54	-5,34	-6,58
Energia van der Waals, ponte de hidrogénio e desolvatação	-5,93	-6,08	-6,3	-5,76	-5,7	-4,77	-6,67	-6,68	-5,0	-6,45
Energia electroestática	-0,66	-0,19	-0,92	-0,32	-0,17	-0,8	-0,38	0,14	-0,34	-0,13
Energia Interna	-1,56	-1,88	-1,24	-2,28	-2,05	-2,1	-1,29	-1,81	-3,17	-1,75
Energia de torção	4,47	4,47	4,47	4,47	4,47	4,47	4,47	4,47	4,47	4,47

No que toca à energia de ligação esta resulta da soma da energia de van der Waals, pontes de hidrogénio, desolvatação, energia electroestática e energia de torção, como foi evidenciado na equação 5. A energia intermolecular resulta da soma do parâmetro da energia de van der Waals, pontes de hidrogénio, desolvatação e energias electrostáticas. Assim a energia de ligação resulta em última análise das contribuições da energia intermolecular com a energia de entropia de torção (parâmetro de penalização). O autodock faz também uma estimativa da constante de inibição (K_i) correspondendo este parâmetro à concentração de ligando necessário para inibir 50% dos locais de ligação na proteína.

II.3.6.2. *Docking* molecular entre AraN-Arabinotriose

Na figura II.38 encontram-se os *dockings* das conformações 6 e 7 respectivamente na estrutura da AraN obtida anteriormente no servido I-TASSER. A conformação 6 foi escolhida visto que é aquela em que a arabinotriose se encontra a interagir com o resíduo W253 conservado na família de SBP's e envolvido no processo de reconhecimento de substratos (figura II.34, mostra que estes resíduo é conservado em todas as proteínas submetidas a alinhamento). A conformação 7 foi escolhida por razões semelhantes visto ser a única na qual a arabinotriose se encontra a interagir com o resíduo R81 da AraN, sendo também este resíduo conservado em algumas SBP's encontrando-se a interagir com o substrato (figura II.35).

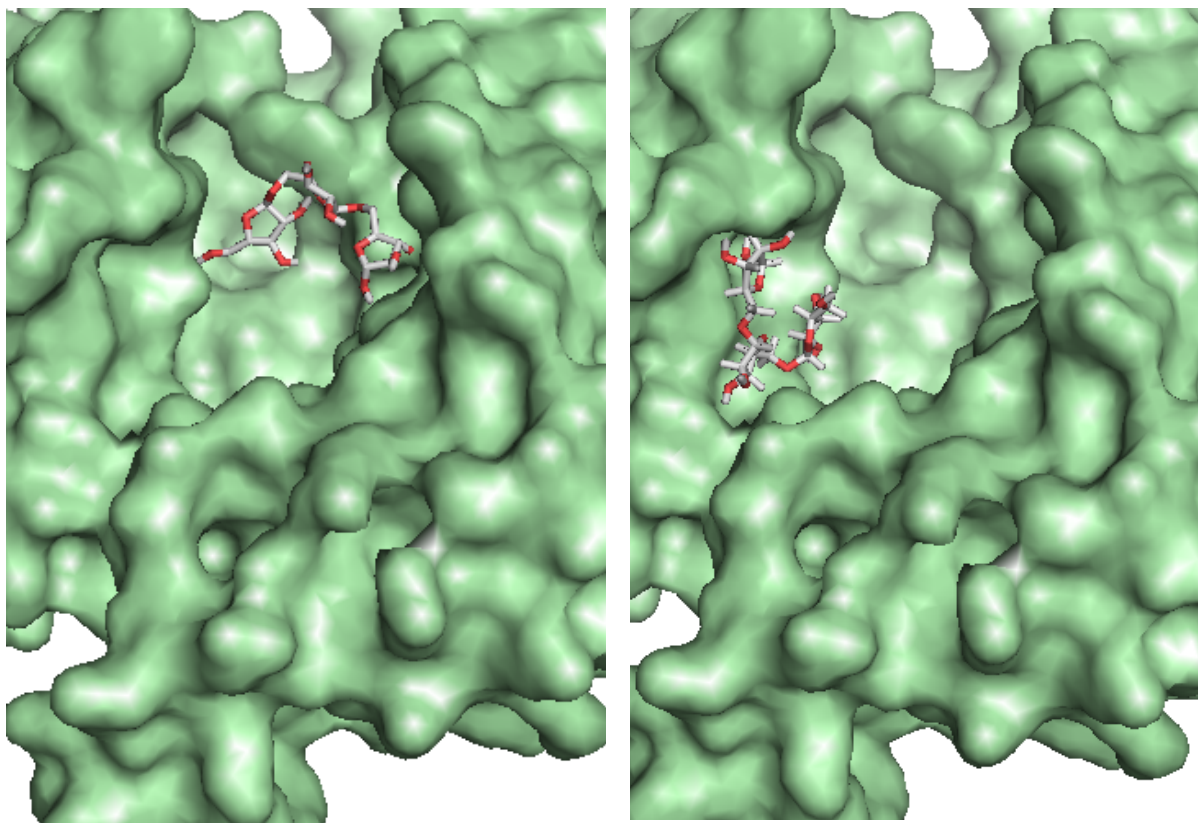


Figura II.38. Conformações obtidas no *docking* molecular. Na figura à esquerda está representado o *docking* AraN-Arabinotriose da conformação 6 e na figura à direita está representado o *docking* da AraN-arabinotriose da conformação 7.

Através da análise de ambos os *dockings* das conformações 6 e 7 (figura II.38), é possível observar que o local de ligação escolhido é distinto. Relativamente, à conformação 6 parece haver uma maior exposição do açúcar ao solvente, ao passo que na conformação 7 a arabinotriose está mais estabilizada numa das “encostas” da AraN. O *docking* com as interações referente à conformação 6 está representado na figura II.39.

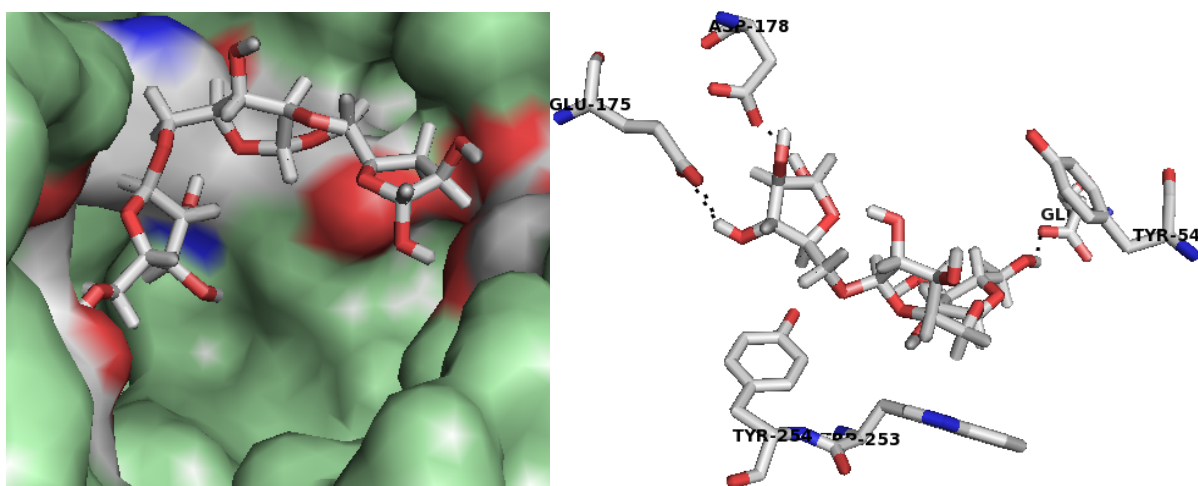


Figura II.39. Mapeamento de interações. Interações AraN-arabinotriose (conformação 6). As figuras têm uma orientação relativa entre elas de 180° .

Na figura II.39 é possível observar que tal como foi previamente observado para outras proteínas a arabinotriose encontra-se estabilizada por três interações hidrofóbicas por CH-*stacking*,

com os resíduos W253, Y254 e Y54. Relativamente ao Y254 este é o resíduo que através da análise da figura II.39 se encontra em maior proximidade com o receptor molecular, estando 0.8 Å do H4 α da subunidade redutora. Este resultado vai ao encontro dos resultados obtidos através do mapeamento de epítomos com recurso ao cálculo do STD_{fit} (II.3.3.1).

Ainda no que diz respeito à subunidade redutora, esta encontra-se envolvida na formação de duas pontes de hidrogénio com a AraN. Estas pontes de hidrogénio são mediadas através do carboxilo da cadeia lateral da E175 com o –OH3, tendo esta ponte de hidrogénio 2.0 Å, bem como pela ponte de hidrogénio (2,0 Å) entre o carboxilo da cadeia lateral do D178 com o –OH1. Este dados estão de acordo com os resultados obtidos por STD-RMN, no qual o protão H3 α da arabinotriose apresenta uma saturação de 35%. A percentagem de saturação do protão H1 α foi de 25%, no entanto a análise deste último protão não pode ser tida em conta visto que o espectro foi obtido através de um sequencia pulso que ao suprimir o sinal residual da água pode irradiar em parte este protão face à sua proximidade do sinal da água.

Os estudos da secção II.3.3.4.3, levantaram a hipótese de na arabinotriose o sinal que maior A_{STD} apresenta é aquele relativo aos protões H2m e H2n. Se tivermos em conta que este mapeamento traduz a proximidade entre a arabinotriose e a AraN, e que um maior afastamento do receptor molecular irá levar em última análise a uma menor transferência de saturação e consequentemente um menor A_{STD}, podemos à partida através da análise da figura II.40 demonstrar que o protão H2m se encontra virado para uma região onde não se encontra a estabelecer contactos com a AraN. Por outro lado o protão H2n fica virado para uma região interna da AraN, nomeadamente para o lado do anel aromático do W253, pese embora esteja a distancia de 3,2 Å da cadeia lateral do resíduo W253. Esta distancia é 4 vezes superior à distancia relativa entre o protão H4 α da subunidade redutora e o resíduo Y254. Assim e de acordo com o modelo obtido por *docking* faria mais sentido que o protão que se encontra-se a receber maior saturação fosse o protão H4 α , algo que pode ser suportado através da análise da figura II.40, que sugere uma maior proximidade da subunidade redutora para com a AraN, fruto das interações que estabelece com o último por pontes de hidrogénio.

Assim pese embora as três subunidades que compõem a arabinotriose apresentem um conjuntos de interações fundamentais no processo de reconhecimento molecular, seria mais espectável que o epítipo de ligação que apresentasse um maior grau de saturação fosse o H4 α , como é explicado na secção II.3.3.1 este facto pode dever-se ao facto do protão H4 α apresentar um T₁ mais curto sendo que o A_{STD} retirado no mapeamento de epítomos, que é efectuado com um tempo de saturação de 1,5 segundos, pode estar a subvalorizar o A_{STD} deste protão.

Ainda relativamente a interação da AraN com a arabinotriose é possível observar a presença de um terceiro resíduo aromático próximo da subunidade não redutora da arabinotriose Y54, sendo que este resíduo é putativo de se encontrar a estabelecer uma interação CH-*stacking*, com um dos CH da subunidade não redutora. O *docking* efectuado não atribuiu liberdade conformacional ao receptor molecular, sendo que numa segunda etapa a execução de *docking* com os resíduos até aqui identificados pode revelar a interação CH-*stacking* entre a Y54 e –CH do subunidade n do açúcar. Ainda nesta subunidade é possível constatar a formação de uma ponte de hidrogénio de 1,8 Å entre o

resíduo E78 e o grupo –OH ligado ao C5 da subunidade final não redutora. A proximidade desta subunidade ao receptor molecular é também ela suportada pelos dados de STD-RMN que revelam que o CH2 que maior saturação recebe por parte da AraN é o da subunidade não redutora com um grau de saturação de 62%.

No que diz respeito a outras SBP's estas reconhecem o açúcar através da sua ligação longitudinalmente ao longo de um local de ligação afilado ao longo da cavidade. Sendo que no caso da AraN a arabinotriose é reconhecida horizontalmente ao longo do local de ligação o que revela que ou por uma lado o local de ligação é mais largo, ou a AraN obtida nos modelos de homologia se encontra na sua conformação aberta. Este último argumento pode ser suportado devido ao facto da proteína estar a estabelecer maioritariamente interações com o domínio N-terminal através dos resíduos E175, D178, W253 e Y254, ao passo que apenas as interações promovidas pelos resíduos E78 e Y54 é que pertencem ao domínio C-terminal [129].

II.3.7. Mutagénese

Como é possível observar na figura II.38 os *dockings* moleculares evidenciaram a possibilidade de processo de reconhecimento molecular ser efectuado em uma de duas regiões opostas da cavidade de ligação na AraN, para tal foram mutados dois resíduos em cada um destes locais para assim se concluir qual o modo de ligação da arabinotriose à AraN. Estes resíduos foram analisados com o auxílio dos alinhamentos e dos *dockings* efectuados.

Um modo de confirmar a importância destes resíduos para o processo de reconhecimento da arabinotriose é através da mutagénese dirigida seguida de um ensaio de STD-RMN para verificar se há alteração no padrão de reconhecimento. Foram então propostas a construção dos mutantes R81A e W253A, encontrando-se este último em progresso.

A sobreexpressão da proteína AraN(R81A) (Mw=47kDa) pode ser confirmada pela presença de uma banda com 47kDa no gel de SDS-PAGE como é possível constatar nos poços de 8 a 9 no gel da figura II.40 e 2 a 10 no gel da figura II.41.

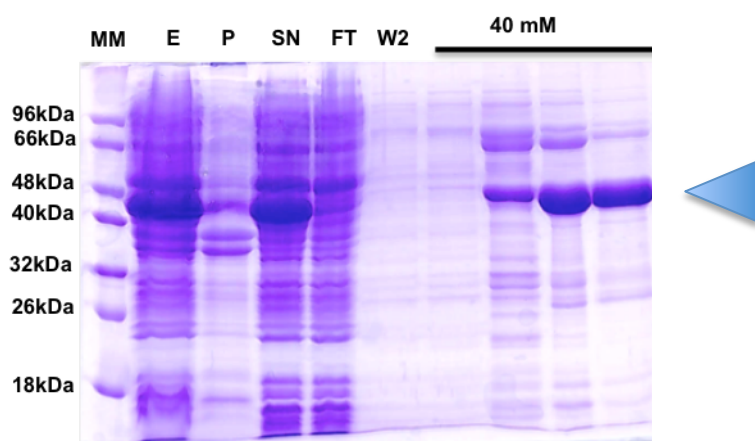


Figura II.40. SDS-PAGE (12,5%), com as fracções de purificação da AraN(R81A). Poço 1- Marcador de pesos moleculares, Poço 2- Extracto, Poço 3- Fração insolúvel (*pellet*), Poço 4- Fração solúvel (sobrenadante), Poço 5- *Flow throw*, Poço 6- *Wash 2*, Poço 7 a 10- Fração eluição 40mM imidazole. Mw(AraN)=47kDa.

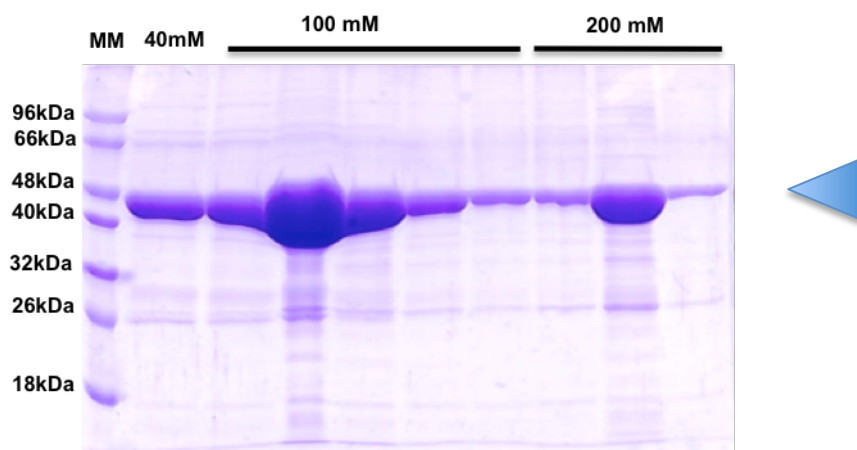


Figura II.41. SDS-PAGE (12,5%), com as fracções de purificação da AraN(R81A). Poço 1- Marcador de pesos moleculares, Poço 2- Fração eluição 40mM. Poço 3 a 7- Fração eluição 100mM imidazole. Poço 8 a 10- Fração eluição 200mM imidazole. Mw(AraN)=47kDa.

Para os estudos de STD foram recolhidas as fracções II, IV, V, VI, VI, IX referentes ao segundo gel (figura II.42), tendo-se obtido 9 mg de AraNR81A pura, por litro de crescimento celular.

Purificada a proteína esta foi testada por STD-RMN em iguais condições aos estudos efectuados previamente: [AraNR81A]= 25 μ M e [Arabinotriose]= 2,5 mM.

Os espectros de STD-RMN obtidos para a AraNwt e AraNR81A na presença de arabinotriose são apresentados na figura II.42 para comparação.

Como se pode ver na figura II.42 a mutação do resíduo R81 não parece afectar a interação entre a AraN e a arabinotriose, pelo que este resíduo não deverá estar envolvido no mecanismo de reconhecimento molecular da arabinotriose.

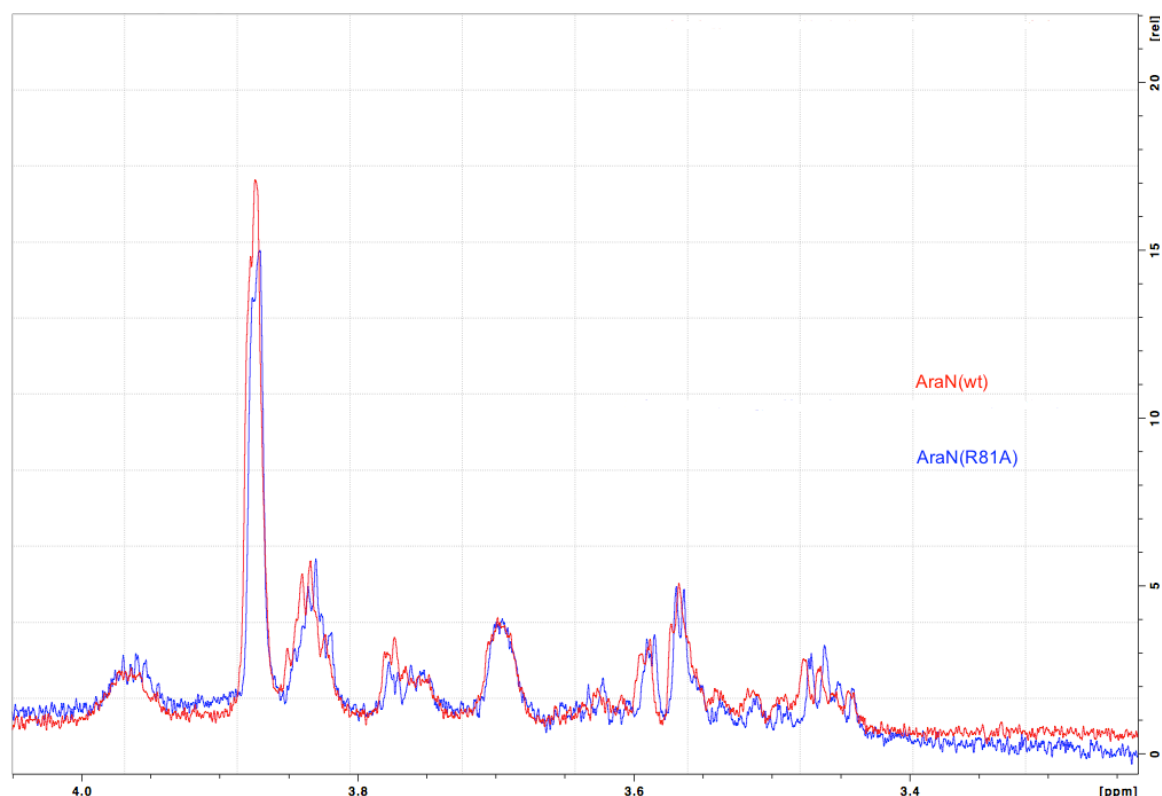


Figura II.42. Espectro de STD-RMN da arabinotriose com AraN(wt) ou AraN(R81). Sobreposição dos espectros ^1H -STD entre a AraN(wt) à concentração de 25 μM na presença de 2,5 mM de arabinotriose e AraN(R81A) à concentração de 25 μM na presença de 2,5 mM de arabinotriose. Ambos os espectros são resultantes da subtração do espectro *off-resonance* (I_0) irradiado a 20000 Hz e do espectro *on-resonance* (I_{SAT}) irradiado a 150 Hz. A aquisição dos espectros foi efectuada a 10°C.

II.3.8. Conformação da arabinotriose no local de ligação da AraN-estudo de NOE de transferência

Uma das primeiras proteínas de reconhecimento de solutos a ser estudada na presença de ligando foi a ABP, na presença de *L*-Arabinose [133]. Esta não evidenciou, uma preferência entre os epímeros α -*L*-Arabinose ou β -*L*-Arabinose. No entanto, no que diz respeito à MBP de *E. coli*, esta apresenta a capacidade de ser estereoespecífica para apenas um dos anómeros, nomeadamente para α -*L*-maltoligossacáridos, sendo este o anómero em menor quantidade em solução [134]. No que diz respeito à interação da arabinobiose com a AraN, quando se considera a subunidade redutora da arabinobiose, verifica-se que o protão H4 α apresenta um grau de saturação de 53%. No entanto, face à sobreposição das ressonâncias espectrais da subunidade redutora - com exceção do protão H4 α - é complicado dizer se há uma preferência conformacional entre a forma anomérica α ou β , ou se apenas ocorre ligação do anómero α -*L*-Arabinobiose.

Através da análise das perturbações dos desvios químicos ao longo da titulação da AraN com o ligando observou-se que a ressonância espectral associada ao protão H1 α desviou 6Hz ao passo que a ressonância espectral associada ao protão H1 β foi perturbada em 3Hz. O facto do protão H1 β sofrer uma perturbação do seu desvio químico é indicativo de que em solução o anómero β de arabinoligómeros também interage com a AraN.

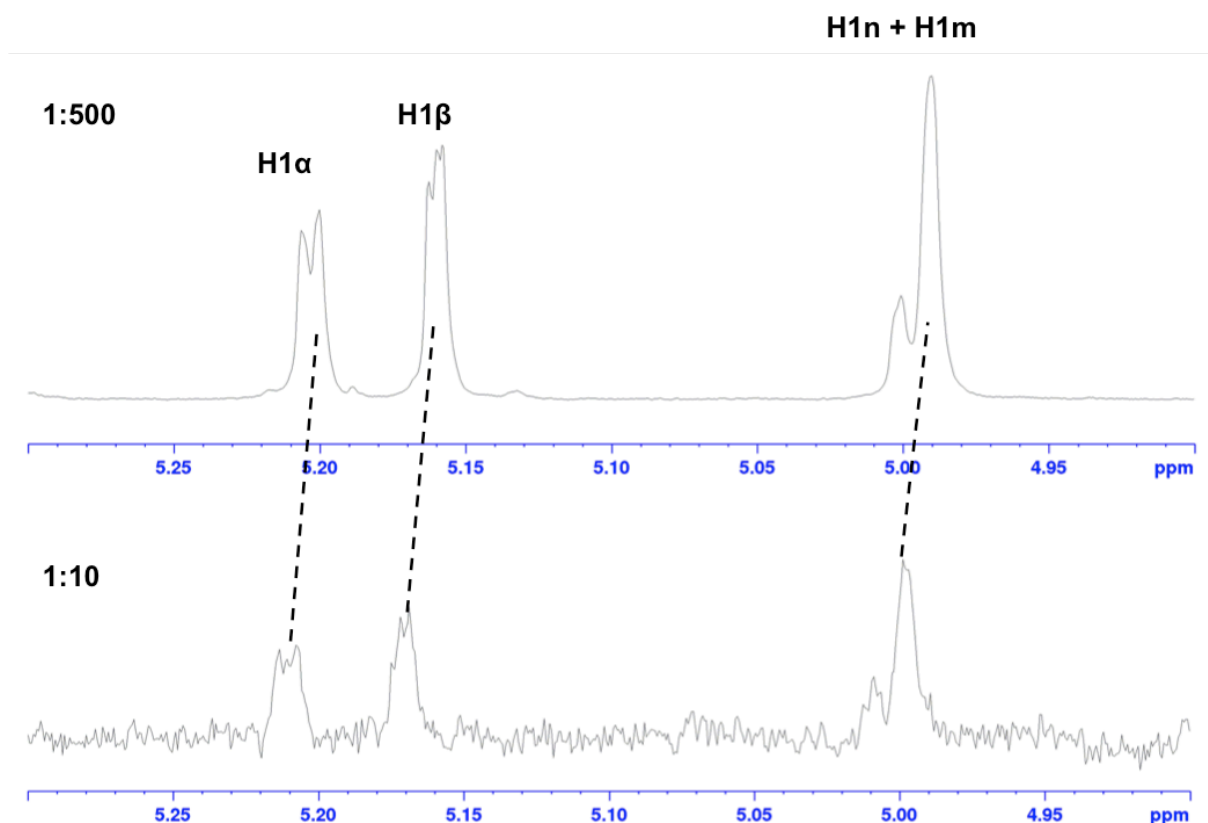


Figura II.43. Perturbação dos desvios químicos ^1H da região anomérica. Ilustração das perturbações dos desvios químicos dos prótons anôméricos da arabinotriose com o aumento da concentração de AraN de 1:10 para 1:500.

Para estudar a conformação do ligando em solução e na local de ligação da proteína, procedeu-se à análise da transferência de NOE entre a proteína e o ligando. Para além do estudo conformacional do carbohidrato, o trNOE dá informação se ocorre ou não ligação entre a proteína e o ligando. Neste último caso o que se verifica é que no caso do ligando interagir com a proteína de maior peso molecular, esta irá adoptar o seu regime de movimento molecular, sendo que as vias de relaxação cruzada adoptadas irão favorecer o desenvolvimento de NOE's negativos. Na conformação livre, os carbohidratos de pequenas dimensões adoptam um regime de NOE positivo, e assim após reconhecimento por parte da proteína é observada uma mudança de fase do sinais que estabelecem as correlações espaciais no espectro de tr-NOESY.

A análise dos espectros nas figuras II.44 e 45, permite tirar duas conclusões: i) ocorre uma clara mudança de fase dos picos cruzados entre o espectro de NOESY e tr-NOESY, confirmando-se que ocorre interação entre a AraN e a arabinotriose. ii) a posição do picos cruzados é mantida inalterada entre o espectro de NOESY e tr-NOESY o que revela que não há uma alteração conformacional associada ao processo de reconhecimento do oligossacárido por parte da AraN. Assim a arabinotriose é reconhecida tal como se encontra em solução.

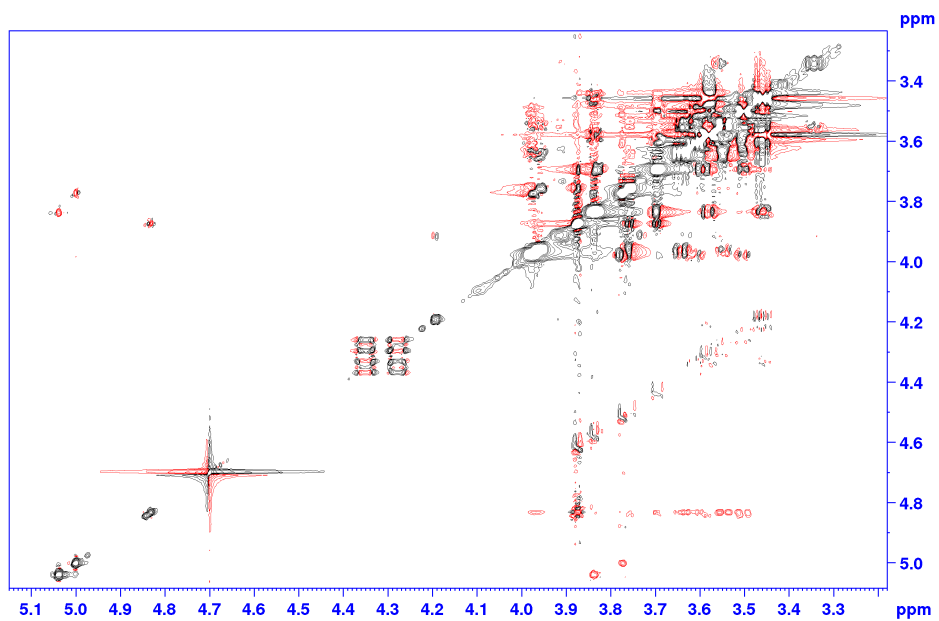


Figura II.44. Espectro 2D-NOESY, da arabinotriose livre sem AraN.

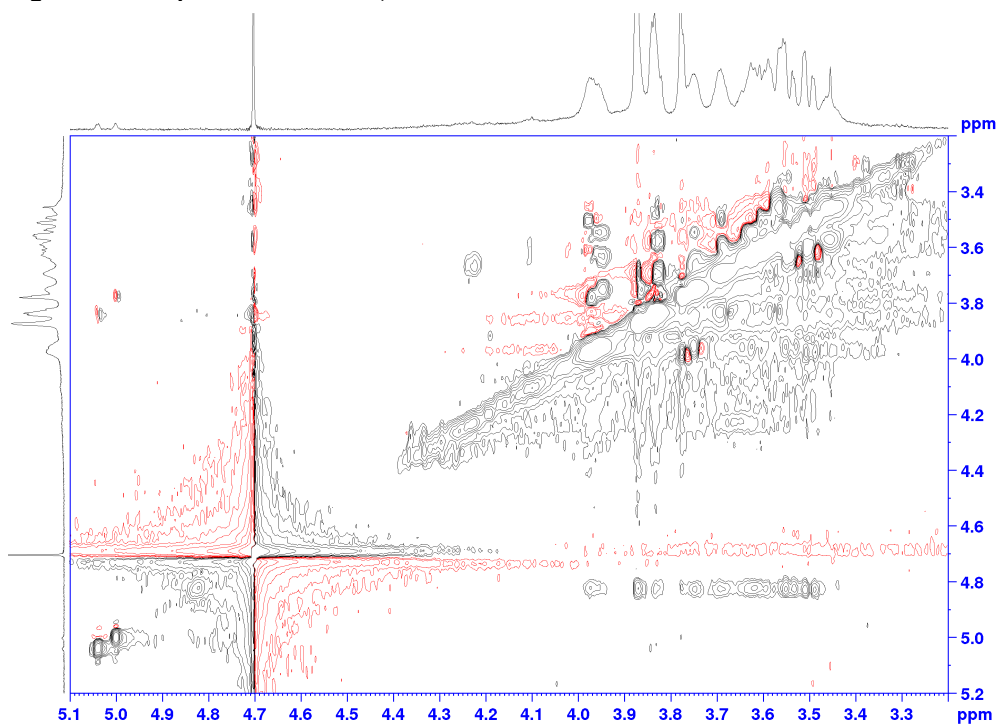


Figura II.45. Espectro 2D-NOESY, da arabinotriose na presença de AraN num excesso de 1:10.

II.4. Conclusões

Através dos estudos efectuados por STD-NMR pode-se concluir que a AraN apresenta capacidade de reconhecer arabinoligossacáridos de 2 até 5 subunidades, não reconhecendo a arabinose enquanto monómero. O facto da AraN conseguir reconhecer a arabinobiose, e os prótons da subunidade não-redutora deste carbohidrato apresentarem o maior grau de saturação revela a importância que a adição desta segunda subunidade têm no processo de reconhecimento molecular. O oligómero de arabinose que apresenta um maior grau de saturação é a arabinotriose o que pode ser indicativo deste ser o tamanho ótimo de reconhecimento no local de ligação.

Com o aumento, consecutivo do tamanho do arabinoligossacárido observa-se que a subunidade central ou procedente à subunidade redutora é aquela que maior grau de saturação recebe, sendo que no caso da arabinotetraose a subunidade não redutora perde saturação comparativamente aos arabinoligossacáridos de 2 e 3 subunidades o que pode suportar a ideia de que a AraN não reconhece uma quarta subunidade deixando-a fora do local de ligação.

No que toca à afinidade do processo de reconhecimento molecular, o K_d têm uma ordem de grandeza de 200 a 700 μM , revelando conseqüentemente uma menor afinidade do que havia sido reportada para outras SBP's na ordem dos 1 a 10 μM .

Para além de reconhecer oligómeros de arabinose, a AraN tem a capacidade de reconhecer oligómeros de glucose na forma de celobiose e celotetraose. Porém, para estas hexaoses na forma de pirano, o reconhecimento dá-se preferencialmente pelas regiões não redutora ao contrario do que fora observado para oligómeros de arabinose. Um facto interessante é o de que a AraN não reconhece a maltohexaose existindo duas possibilidades para justificarem este facto: i) o mecanismo de reconhecimento pode ser estereoespecífico, ou seja, a AraN apenas reconhece o epímero β dos monómeros de β -glucose que compõem os celoligossacáridos, não reconhecendo o epímero α que compõem os α -1,4-maltoligossacáridos, ou ii) o facto da maltohexaose ser composta por seis subunidades e o local de ligação reconhecer otimamente três subunidades ficando as restantes mais expostas ao solvente, podem ser uma razão pela qual na competição a arabinotriose apresentar um maior A_{STD} subvalorizando a interação da maltohexaose.

Estudos de STD em condições de competição sugerem que a constante de afinidade que caracteriza o processo de reconhecimento da AraN para com arabinoligossacáridos e celoligossacáridos, tem uma ordem de grandeza semelhante, na competição pelo local de ligação.

Através de estudo *in silico* foi possível identificar resíduos envolvidos na interação AraN-arabinotriose, maioritariamente de natureza hidrofóbica através do estabelecimento de interações CH-*stacking* através dos resíduos W253, Y254 e Y54, bem como o estabelecimento de pontes de hidrogénio com a subunidade redutora através dos resíduos E175 e D178, e entre a subunidade não redutora e o resíduo E78. Nesta conformação o próton H4 α da arabinotriose é o que apresenta uma maior proximidade para a AraN. Através de estudos de mutagénese e STD-RMN foi possível concluir que o resíduo R81 não desempenha um papel essencial no reconhecimento de arabinoligossacáridos, visto que o espectro de STD-RMN efectuado com o a AraN(wt) e AraN(R81A) deu origem a um espectro de STD igual, sendo que daí se conclui que apenas a mutação do resíduo R81 não irá produzir diferenças significativas no processo de reconhecimento molecular.

Finalmente através de estudos de NOESY na presença e ausência de proteína verificou-se que a arabinotriose é reconhecida por parte da AraN. A conformação através da qual a arabinotriose é reconhecido é igual à sua conformação livre em solução.

Capítulo III: Otimização da sobreexpressão da AraL

III.3.1. Introdução

A AraL é um membro bacteriano da superfamília ubíqua das desalogenases de haloalcanos (DAH), atribuída à família IIA das DAH's. Esta enzima exibe baixa especificidade e atividade catalítica para vários carboidratos fosforilados. Com o intuito de elucidar e caracterizar o processo de reconhecimento irão ser utilizadas metodologias de sobreexpressão e solubilização/renaturação da mesma para subsequente estudo de interação enzima-substrato do ponto de vista da proteína. As metodologias a utilizar encontram-se resumidas na figura III.1.

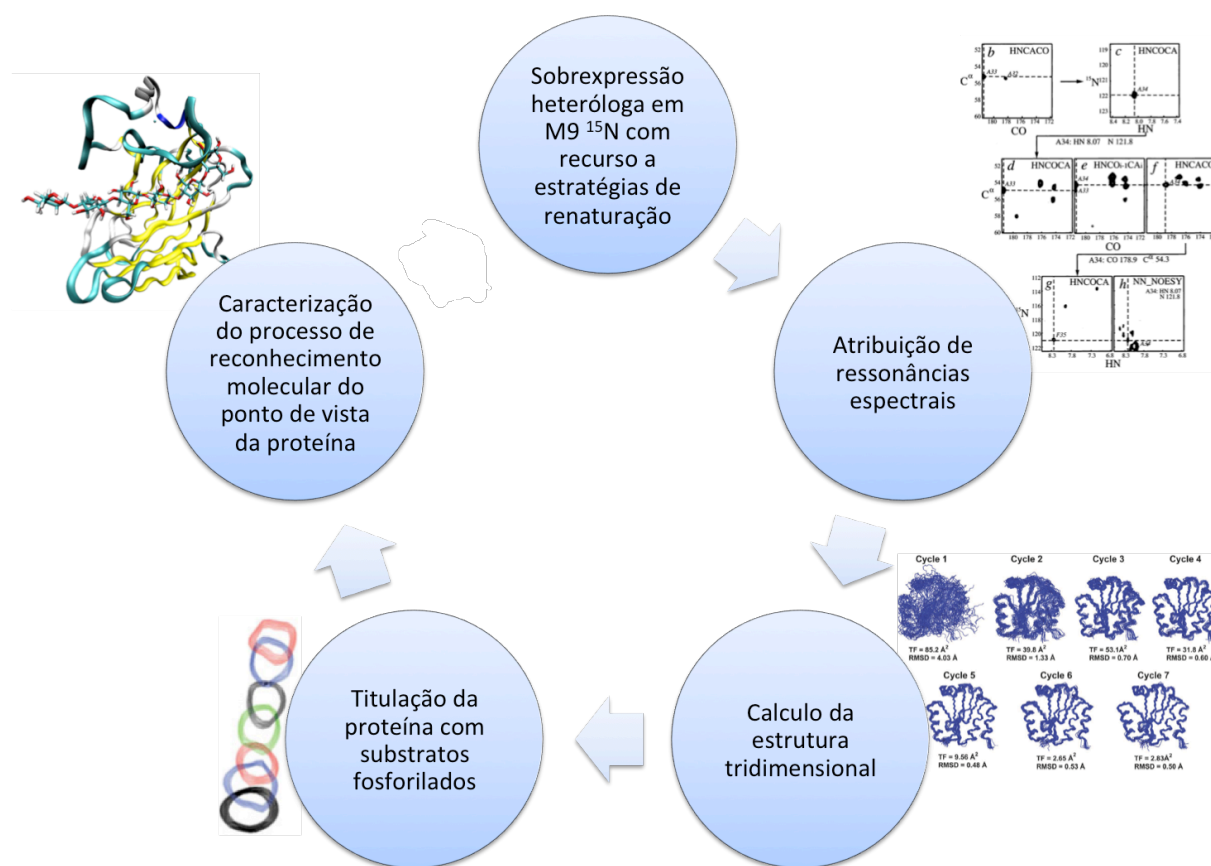


Figura III.1. Estratégia utilizada para estudar a interação protein-carboidrato promovida por parte da AraL. Inicialmente a proteína é sobreexpressa em meio M9 suplementado com ^{15}N sendo solubilizada e renaturada. Seguidamente é necessário fazer as atribuições espectrais da mesma, bem como calcular a sua estrutura tri-dimensional. Em último lugar esta pode ser titulada com um carboidrato fosforilado, para validação de i) mecanismos de reconhecimento molecular, ii) domínios e resíduos envolvidos neste processo.

III.3.2. Materiais e métodos

III.3.2.1. Sobreexpressão da AraL-His₆Tag

A expressão da AraL-His₆tag foi efectuada a partir de um *stock* de glicerol de células competentes de *E.coli* BL21(DE3) pLysS (Novagen), transformadas com o vector pLG11, um derivado de pET30a (Novagen). Este vector, contem a região codificante do gene *araL* sob o controlo do promotor T7, e permite a expressão de uma forma recombinante de AraL que possui uma cauda de histidina na região C-terminal [10].

A proteína foi sobreexpressa em meio LB com os antibióticos Cm (25 µg.mL⁻¹) e Km (20 µg.mL⁻¹). Para efeitos de teste de sobreexpressão (secção III.3.2.2.1) e *screening* de condições de solubilização de corpos de inclusão (secção III.3.2.2.2) foram efectuados crescimentos de 20 mL. Para os estudos de solubilização de corpos de inclusão com recurso a SDS (secção III.3.2.2.3), ureia (secção III.3.2.2.4), GdmCl (secção III.3.2.2.5) e renaturação da AraL foram efectuados crescimentos de 1L para o primeiro e 200 mL para os últimos (ureia e GdmCl). Em todos os crescimentos celulares, as células são crescidas a 37°C até a OD₆₀₀=0,6, procedendo-se à adição de ITPG à concentração final de 1 mM. O tempo de indução é de 4 horas, a única exceção foi na secção III.3.2.2.1 em que se testou um tempo de indução de 3 horas e 30 minutos e O/N (*overnight*).

III.3.2.2. Isolamento e purificação da AraL-His₆Tag

Tendo em conta que a AraL forma corpos de inclusão foram utilizadas diferentes estratégias de solubilização e renaturação da mesma utilizando por um lado agentes caotrópicos como a ureia (6M), guanidino hidrocloreídrico (GdmCl) (6M) e detergentes como é o caso do SDS (1%).

III.3.2.2.1. Teste de sobreexpressão à escala piloto

Foi efectuado um crescimento de 20 mL a 37°C até à OD₆₀₀=0,6 sendo esse crescimento subdividido em 2 crescimentos cada um composto por 10 mL, sendo que um dos crescimentos não é induzido e o segundo é induzido, com o intuito de confirmar a sobreexpressão da AraL. No que toca ao crescimento induzido, são recolhidas duas alíquotas uma ao fim de 3 horas e 30 minutos de indução e a segunda recolha é efectuada O/N (*overnight*).

Subsequentemente as células são recolhidas por centrifugação a 4°C durante 10 minutos (Sigma 3-16K Sartorius/ Rotor:19776) a 7500 rpm.

De seguida as células são partidas com auxílio do sonicador (Hielscher UP50H/UP100H ponta MS7) em ciclos de amplitude 70%, durante 30 segundos, com 30 segundos de intervalo entre cada ciclo, sendo este processo repetido durante 7 ciclos. Para se isolar as proteínas é efectuada uma centrifugação durante 15 minutos a 9500 rpm's. Finalmente procede-se ao SDS-PAGE (sendo a percentagem de acrilamida de 12,5%) destes extratos (figura III.3).

III.3.2.2.2. Screening de condições de sobreexpressão

Com recurso a um protocolo EMBL (Protein Purification- Extraction and Clarification- Solubility Studies- EMBL- http://www.embl.de/pepcore/pepcore_services/protein_purification/extraction_clarification/solubility_studies/) , foram testadas diferentes condições de solubilização e

purificação com o intuito de validar de um modo rápido e eficiente se algumas das condições permitiriam a obtenção de AraL solúvel. Para tal foi efectuado um crescimento de 20 mL a 37°C até à $OD_{600}=0,6$, e subsequentemente adicionado IPTG à concentração final de 1 mM, com um tempo de indução de 4 horas. As células foram recolhidas por centrifugação a 4°C durante 10 minutos a 7500 rpm. Subsequentemente as células foram resuspensas em tampão e partidas com auxílio do sonicador em ciclos de amplitude 70%, durante 30 segundos, com 30 segundos de intervalo entre cada ciclo, sendo este processo repetido durante 7 ciclos. O isolamento das proteínas é efectuado com recurso à centrifugação durante 15 minutos a 9500 rpm's.

Foram testados diferentes tampões para resuspender as células: i) 50mM Tris pH=7,5, 50mM NaCl, 5mM EDTA, 1mg/mL lisozima (tampão 7,5N) ii) 50mM Tris pH=7,5, 2M NaCl, 5mM EDTA, 1mg/mL lisozima (tampão 2S), 50mM Tris pH=7,5, 50mM NaCl, 5mM EDTA, 0,5M ureia, 1 mg/mL lisozima (tampão 0,5U), iv) 20mM Tris pH=7,5, 50mM NaCl, 0,2 % Tween 20, 1mg/mL lisozima (tampão T).

Finalmente procede-se ao SDS-PAGE (sendo a percentagem de acrilamida de 12,5%) destes extratos (figura III.4 e figura III.5).

III.3.2.2.3. Solubilização de corpos de inclusão com ureia e renaturação na coluna Ni-NTA

As células são recolhidas por centrifugação a 4°C durante 10 minutos a 7500 rpm. A fração insolúvel é resuspensa em tampão de ligação (5mM imidazole, 500mM NaCl, 20mM Tris-HCl pH=7,5, 1mM β -Mercaptoetanol, 15mM $MgCl_2$ e 25 μ g/mL PMSF). De seguida as células são partidas com auxílio do sonicador em ciclos de amplitude 70%, durante 30 segundos, com 30 segundos de intervalo entre cada ciclo, sendo este processo repetido durante 7 ciclos.

A AraL é recolhida por centrifugação a 10 000g durante 30 min a 4°C, sobre a forma de corpos de inclusão na fração insolúvel. Esta fração é consequentemente resuspensa em tampão de lavagem (5mM imidazole, 500mM de NaCl, 20mM Tris-HCl pH=7,5, 1mM β -Mercaptoetanol, 15mM $MgCl_2$ e 1% Triton X-100), sendo lavada duplamente com a recolha do *pellet* através da centrifugação a 4°C durante 30 minutos a 10 000g. De seguida a fração proteica insolúvel foi resuspensa em tampão de equilíbrio (5mM imidazole, 500mM de NaCl, 20mM de Tris-HCl pH=7,5, 15mM de $MgCl_2$, 1mM β -Mercaptoetanol, 6M Ureia) [135, 136]. Após a proteína ser solubilizada a fração é centrifugada e o sobrenadante filtrado com uma membrana de 0,45 μ m.

Após a coluna ter sido funcionalizada com níquel e equilibrada com 10 volumes de tampão de equilíbrio a coluna é carregada com o sobrenadante previamente filtrado. De seguida é efectuada a renaturação da proteína dentro da coluna, através da redução gradual da força iónica de ureia até a coluna ficar sem ureia (6M ureia a 0M de ureia) [137]. A redução gradual de concentração de ureia é feita através da passagem de 5 frações de tampão de renaturação (5mM imidazole, 500mM de NaCl, 20mM de Tris-HCl pH=7,5, 1mM β -Mercaptoetanol, 15mM $MgCl_2$, e ureia em quantidades decrescentes de 6M, 5M, 4M 3M, 2M, 1M, 0M). Finalmente e como foi descrito para a AraN a AraL é eluída com recurso a um gradiente de imidazole.

III.3.2.2.4. Solubilização de corpos de inclusão com guanidino de hidrocloreto e renaturação na coluna Ni-NTA

Este protocolo é análogo ao protocolo de solubilização e purificação com ureia, sendo que para os efeitos deste protocolo o agente caotrópico utilizado foi o GmdCl (guanidino hidrocloreto) a uma concentração de 6M. Os paços de isolamento e purificação foram iguais ao método anterior.

III.3.2.2.5. Solubilização de corpos de inclusão com SDS

As células são recolhidas por centrifugação a 4°C durante 10 minutos a 7500 rpm. O *pellet* foi resuspenso em tampão de ligação (8mM Na₂HPO₄, 286mM NaCl, 1,4 mM KH₂PO₄, 2,6 mM KCl, 1%SDS (p/v), 1mM β-Mercaptoetanol a pH=7,4). De seguida as células são partidas com auxílio do sonicador em ciclos de amplitude 70%, durante 30 segundos, com 30 segundos de intervalo entre cada ciclo, sendo este processo repetido durante 7 ciclos [138].

De seguida é adicionado ao lisado KH₂PO₄ à concentração final e 150mM e deixado incubar em gelo durante 30 minutos (ou até se observar o dodecil sulfato precipitado) [139]. Após este passo a mistura é centrifugada a 13 000 rpm durante 30 minutos a 4°C. O *pellet* resultante é lavado num tampão de lavagem (8mM Na₂HPO₄, 286mM NaCl, 1,4mM KH₂PO₄, 2,6mM KCl, 0,1% Sarkosyl (p/v), 1mM β-Mercaptoetanol a pH=8,5). Este sobrenadante é novamente incubado em gelo durante 30 minutos. Seguidamente é efectuada a centrifugação a 13000rpm durante 30 minutos a 4°C, sendo que o sobrenadante é recolhido e filtrado num filtro de 0,45µM, antes de ser carregado na coluna de afinidade Ni-NTA [139].

Após a coluna ter sido funcionalizada com níquel esta é equilibrada com o tampão de equilíbrio (8mM Na₂HPO₄, 286mM NaCl, 1,4 mM KH₂PO₄, 2,6 mM KCl, 0,1% de Sarkosyl (p/v), 1mM β-Mercaptoetanol) e carregada com a proteína, sendo posteriormente lavada com o tampão de equilíbrio cinco vezes com o intuito de remover quantidades residuais de SDS. Finalmente a proteína é eluída com recurso a um gradiente de imidazole (40mM, 100mM, 200mM e 500mM de imidazole).

III.3.2.3. Verificação de atividade fosfatase da AraL

Para aferir qualitativamente a recuperação de atividade por parte de AraL após renaturação, foi utilizado pNPP (para-nitrofenil fosfato) como substrato. A mistura reacional é composta por AraL (20µg) em tampão 100mM Tris-HCl a pH=7, com 15mM MgCl₂, o substrato (pNPP) encontra-se à concentração 20mM pNPP. Esta mistura é incubada a 65°C durante 1 hora, sendo que no final a reação é parada pela adição de 1mL de NaOH à concentração final de 0,2M. Os ensaios foram efectuados com diluição da AraL 1:1, 1:2 e 1:4 sendo as absorvâncias medidas em triplicado. A atividade é medida através da leitura da absorvância a 405nm, contra um padrão semelhante à amostra em questão mas desta feita sem AraL. Após a AraL ser purificada e antes de se efetuarem as medições de atividade esta é dialisada contra o tampão em cima indicado durante 48 horas.

III.3.2.4. Monitorização do processo de solubilização e renaturação da AraL com recurso a SDS por RMN

Todos os espectros foram adquiridos no espectrofotómetro de 400 MHz Bruker Avance III, equipado com uma sonda TXI, operando a uma frequência de 400.15 MHz em ^1H . Os espectros ^1H foram adquiridos com a sequência de pulsos zgpg30, ao final de 16 scans numa janela espectral de 6410,3 Hz (16 ppm).

III.3.3. Resultados e discussão

III.3.3.1. Solubilização e renaturação de corpos de inclusão

Com o desenvolvimento de metodologias de sobreexpressão de proteínas através de DNA recombinante, abriu-se uma janela de oportunidades ao nível da área da biologia estrutural e funcional. O uso de sistemas bacterianos para sobreexpressão de proteínas permite a obtenção de elevadas quantidades das mesmas, pese embora em muito casos estas proteínas face aos elevados níveis de sobreexpressão sofram processos de agregação e acumulação sobre a forma de corpos de inclusão disfuncionais. Com o intuito de estudar as funções biológicas destas proteínas é necessário que estas se encontrem no seu estado nativo na sua conformação ativa.

A sobreexpressão de proteínas sobre a forma de corpos de inclusão pode ser um verdadeiro paraíso ou uma autentica dor de cabeça. No primeiro caso, e no que toca à sobreexpressão de proteínas à larga escala para fins farmacêuticos, estas irão acarretar um menor gasto económico nos processos de purificação a jusante, visto que os corpos de inclusão irão ser constituídos exclusivamente da proteína sobreexpressa (potenciais contaminantes podem estar na forma de proteínas membranares que face à sua hidrofobicidade irão ficar retidas no *pellet* em conjunto com os corpos de inclusão embora não façam parte dos mesmos) que em alguns casos representa cerca de 50% do material protéico da células em questão. Face à elevada densidade dos corpos de inclusão estes podem ser facilmente isolados com recurso à sedimentação dos mesmos por centrifugação [140]. Outra vantagem da obtenção de proteína recombinante sobre a forma de corpos de inclusão é sua maior proteção contra a degradação proteolítica.

No entanto, a solubilização destes agregados proteicos e consequente renaturação podem ser bastante problemáticos, na medida que o processo cinético que caracteriza o processo de enrolamento de proteínas é bastante difícil de mimetizar. Ao longo do processo de renaturação, pode por um lado voltar a ocorrer agregação de proteínas, quando esta se encontra num estado intermédio entre o estado desnaturado e nativo, e por outro um incorreto estabelecimento de interações ao nível da formação de motivos estruturais ao nível da estrutura terciária, visto que o processo de renaturação não é o único a ocorrer competindo com processos de agregação e *misfolding* [141]. Assim a taxa (rapidez) com que ocorre o processo de renaturação é determinante para a correta renaturação.

O processo de renaturação leva à modificação conformacional da proteína do seu estado desenrolado para o estado nativo enrolado. Na presença de elevadas concentrações de agente desnaturante, as proteínas irão encontrar-se altamente desordenadas, flexíveis e solvatadas. Num solvente polar como a água estas irão encontrar-se mais compactas e enroladas. Assim o processo de renaturação com recurso a agentes caotrópicos tem como objectivo mimetizar este processo. Idealmente a transferência de uma proteína num meio desnaturante para um meio aquoso irá propulsionar o processo de enrolamento com a proteína obrigando-a a colapsar numa estrutura compacta e organizada. No entanto se este processo for drástico podem ocorrer formações de agregados ou pode levar ao enrolamento incorreto da mesma. O segredo para um correto enrolamento reside na concentração intermédia de agentes desnaturante ao longo da redução

gradual do mesmo, devendo esta concentração permitir por um lado o colapsamento da proteína para uma estrutura rígida, e por outro permitir um determinado grau de flexibilidade à proteína para que esta se rearranje conformacionalmente durante o processo de renaturação [141].



Figura III.2. Landscape dos estados de enrolamento da proteína em função da concentração de agente desnaturante utilizado.

Quando o processo de renaturação é problemático a formação de corpos de inclusão pode ser evitada através de diversas alternativas: i) tendo em conta que o rendimento de proteína nativa aumenta com a diminuição da taxa de expressão de proteínas recombinantes, uma alternativa passa pela sobreexpressão das mesmas a uma taxa mais baixa pela diminuição da temperatura da cultivo celular [142], ii) utilização de sistemas de co-expressão com chaperonas, iii) utilização Tag's de solubilidade acopladas à proteína recombinante, caso o motivo pelo qual estas se deposite nos corpos de inclusão esteja relacionado com o facto de esta apresentar regiões hidrofóbicas à sua superfície [143].

No que diz respeito à obtenção da AraL na sua forma nativa foram utilizados procedimentos de solubilização de corpos de inclusão através de Ureia, GdmCl ou SDS, na presença de agentes redutores tais como o β -Mercaptoetanol, visto que uma dos veículos condutores à formação de corpos de inclusão está intimamente correlacionado com a formação de pontes de dissulfídicas entre proteínas, num agregado sob a forma de corpo de inclusão.

III.3.3.1.1. Teste de sobreexpressão da AraL

Pela análise do gel de SDS-PAGE na figura III.3. é possível observar a presença de uma banda carregada, nomeadamente no poço 4 e poço 6, próxima dos 32 kDa que corresponde à AraL (Mw=29 kDa). A observação da sobreexpressão da AraL fica ainda mais evidente quando comparada com as fracções 2 e 3, na qual a sua sobreexpressão não é induzida, que corresponde ao poço 2 e 3 da figura III.3. Deste teste de sobreexpressão surge também a informação de que na sua grande maioria esta proteína irá encontrar-se na fracção insolúvel (poço 4 e 6 representam as fracções induzidas insolúveis, ao passo que as fracções presentes no poço 5 e 7 são as fracções induzidas solúveis).

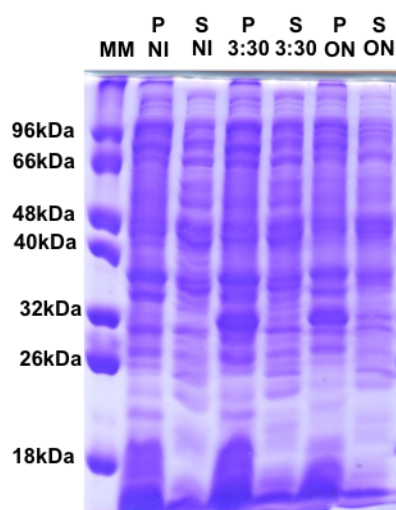


Figura III.3. SDS-PAGE (12,5%), com as fracções de purificação da AraL. Poço 1- Marcador de pesos moleculares, Poço 2- Fração insolúvel não induzida, Poço 3- Fração solúvel não induzida, Poço 4- Fração insolúvel 3 horas e 30 minutos de indução, Poço 5- Fração solúvel 3 horas e 30 minutos de indução, Poço 6- Fração insolúvel indução O/N, Poço 7- Fração solúvel com indução O/N. Mw(AraL)=29 kDa.

A análise qualitativa da sobreexpressão da proteína AraL não revela grande diferença nos rendimentos de sobreexpressão se esta for sobre-expressa durante 3 horas e 30 minutos e ao longo da noite (O/N).

III.3.3.1.2. *Screening* de tampões de solubilização da AraL

Numa primeira fase deste projeto, e tendo em conta que um dos principais objectivos seria o de hipoteticamente determinar a estrutura tri-dimensional da proteína era necessário, de um modo rápido e eficaz encontrar condições que permitissem solubilizar e renaturar a AraL, sempre com elevados rendimentos em vista.

Para tal numa primeira etapa foram estudados protocolos de *screening* de tampões de solubilização da AraL. Na figura III.4 e figura III.5 encontram-se esses mesmos resultados. Em nenhum dos casos se verificou um aumento significativo de solubilização da AraL.

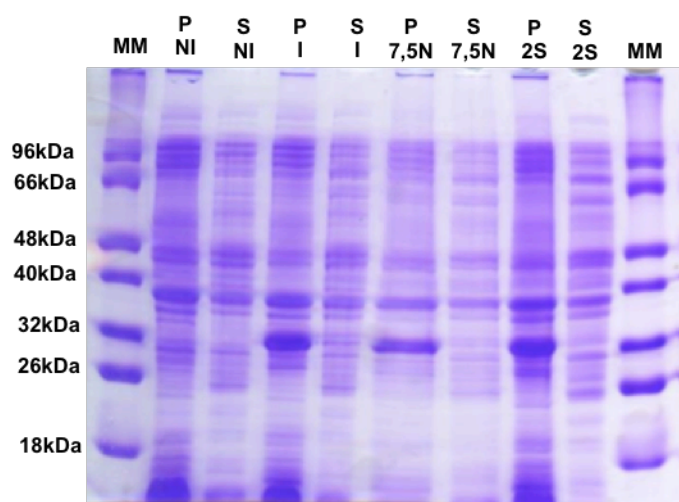


Figura III.4. SDS-PAGE (12,5%), com as fracções de purificação da AraL. Poço 1- Marcador de pesos moleculares, Poço 2- Fração insolúvel não induzida, Poço 3- Fração solúvel não induzida, Poço 4- Fração insolúvel resuspensa em tampão (20 mM PBS pH=7,4, 62,5 mM NaCl, 10mM imidazole), Poço 5- Fração solúvel

resuspensão em tampão (20 mM PBS pH=7,4, 62,5 mM NaCl, 10mM imidazole), Poço 6- Fração insolúvel após tratamento com tampão 7,5N, Poço 7- Fração solúvel após tratamento com tampão 7,5N, Poço 8- Fração insolúvel após tratamento com tampão 2S, Poço 9- Fração solúvel após tratamento com tampão 2S. Poço 10- Marcador de pesos moleculares.

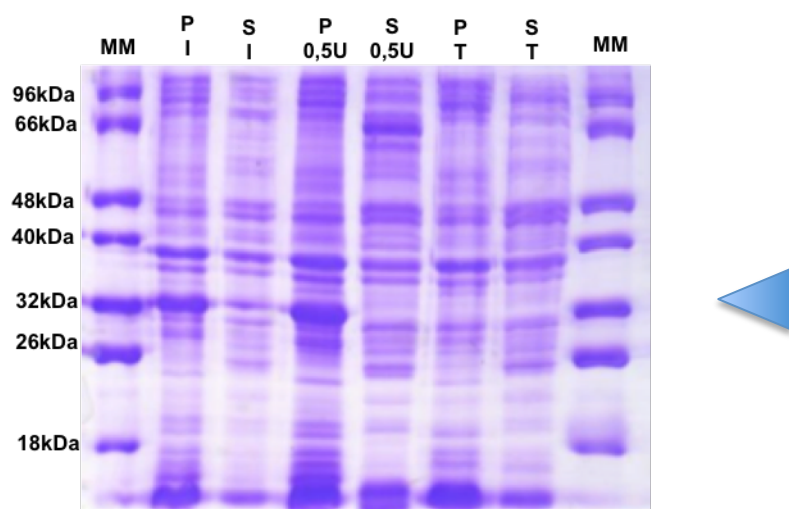


Figura III.5. SDS-PAGE (12,5%), com as fracções de purificação da AraL. Poço 1- Marcador de pesos moleculares, Poço 2- Fração insolúvel induzida, Poço 3- Fração solúvel induzida, Poço 4- Fração insolúvel após tratamento com tampão 0,5 U, Poço 5- Fração solúvel após tratamento com tampão 0,5 U, Poço 6- Fração insolúvel após tratamento com tampão T, Poço 7- Fração solúvel após tratamento com tampão T. Poço 8 - Marcador de pesos moleculares.

III.3.3.1.3. Solubilização e renaturação da AraL com recurso a SDS

Numa primeira fase à escala piloto foi testada a eficiência de remoção do SDS com recurso ao 180 mM de KH_2PO_4 e 100 mM de KCl [139]. Após adicionados os sais, a mistura reacional contendo proteína e 1% (p/v) de SDS ficou imediatamente turva, tendo estas amostras sido incubadas em gelo. Seguidamente as amostras foram centrifugadas a 13 000 rpm's durante 5 minutos, sem arrefecimento de centrifuga.

Na figura III.6 e III.7 encontram-se os espectros ^1H dos sobrenadantes sem remoção de SDS, e com remoção de SDS com recurso à adição de KH_2PO_4 ou KCl.

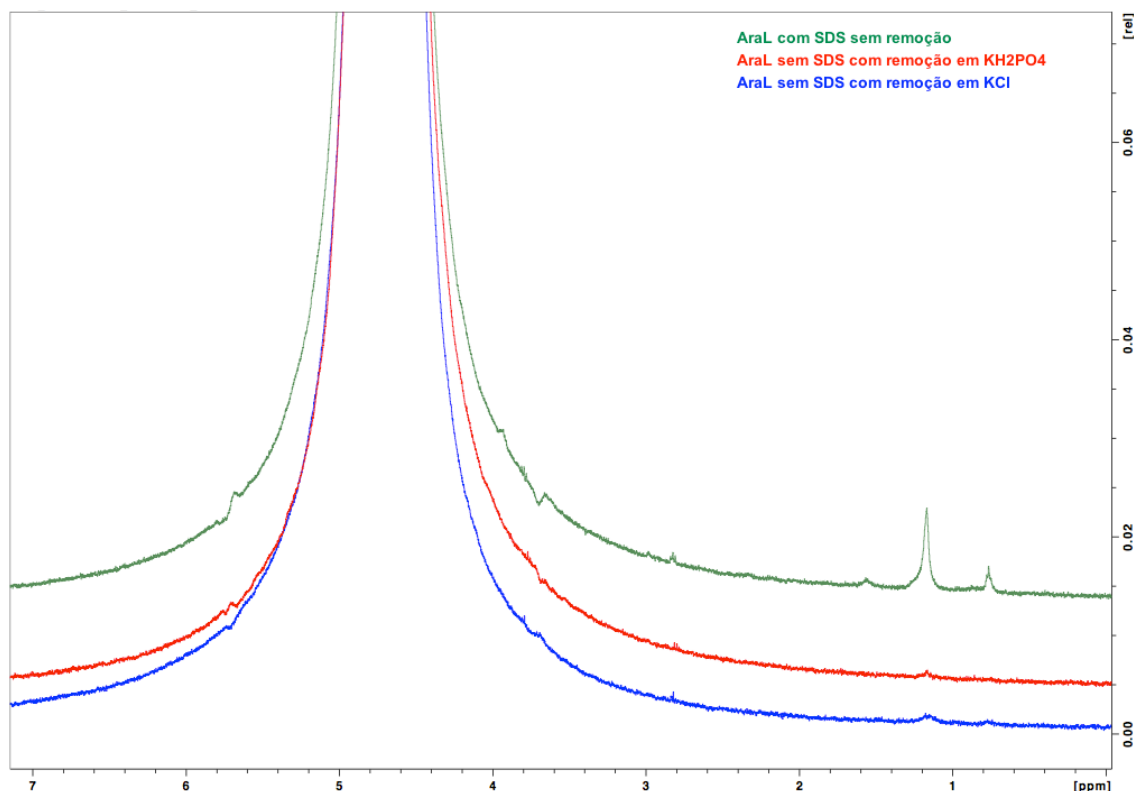


Figura III.6. Espectro ^1H dos diferentes sobrenadantes. A verde encontra-se o espectro da fracção do sobrenadante sem remoção do SDS, a vermelho encontra-se o espectro do sobrenadante com remoção do SDS com recurso a KH_2PO_4 e finalmente a azul encontra-se o espectro do sobrenadante com a precipitação do SDS a ser efectuada pela adição de KCl conseguidos através das diferentes estratégias de remoção de SDS.

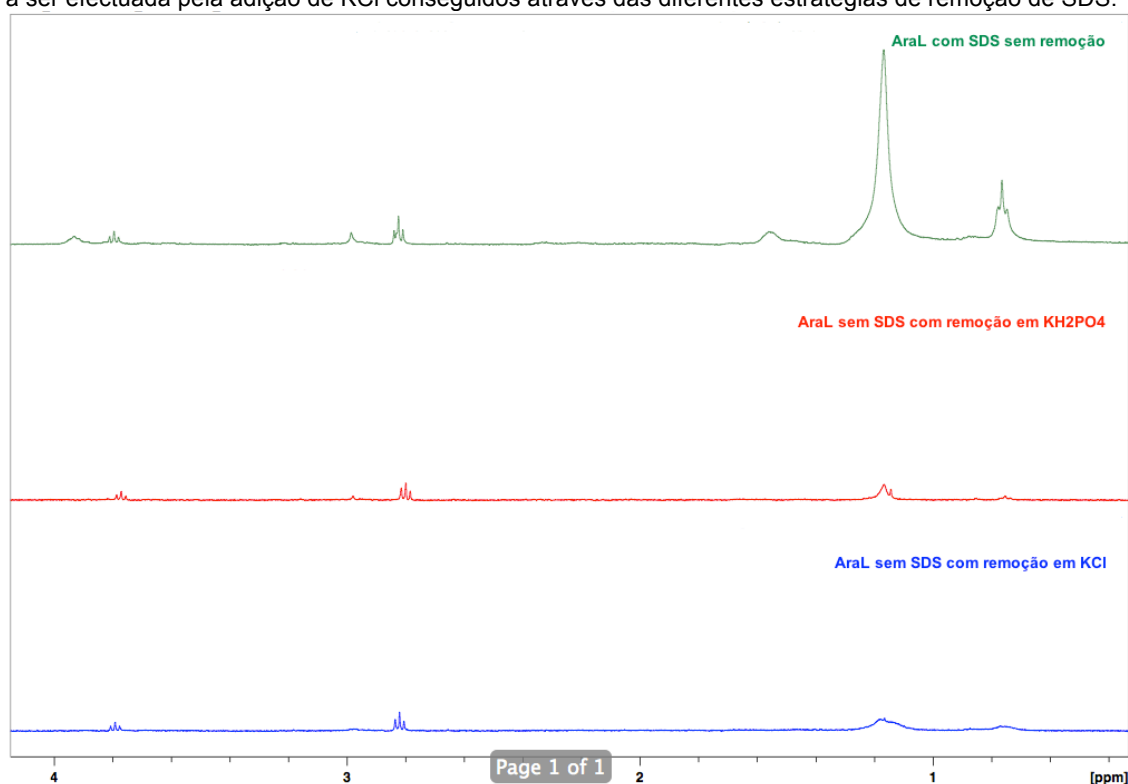


Figura III.7. Espectro ^1H (com supressão de água- zgesgp) dos diferentes sobrenadantes. A verde encontra-se o espectro da fracção do sobrenadante sem remoção do SDS, a vermelho encontra-se o espectro do sobrenadante com remoção do SDS com recurso a KH_2PO_4 e finalmente a azul encontra-se o espectro do sobrenadante com a precipitação do SDS a ser efectuada pela adição de KCl conseguidos através das diferentes estratégias de remoção de SDS.

A região espectral presente na figura III.6. mostra que há uma putativa redução da intensidade das ressonâncias espectrais dos metilos da cadeia alifática do dodecil sulfato (anião). Na figura III.8 e III.9 é possível confirmar a capacidade solubilizadora do SDS em grande extensão. A presença da AraL pode ser confirmada nas frações de eluição 8 a 10 da figura III.7.

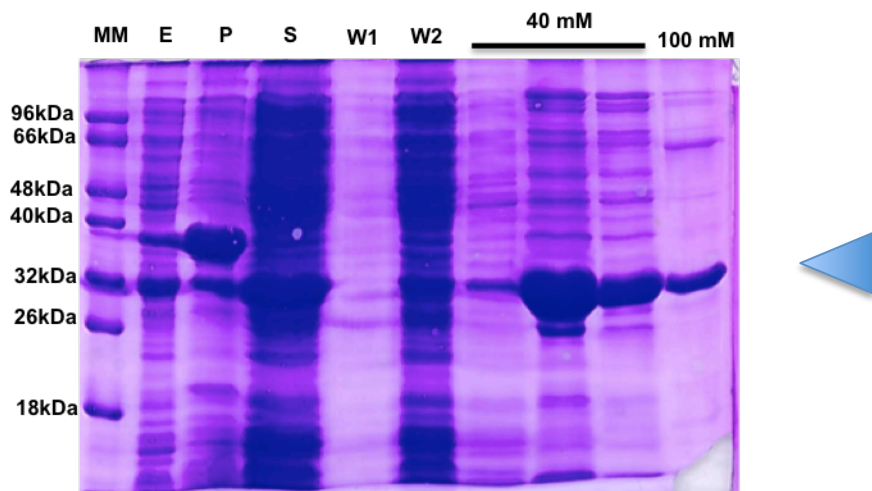


Figura III.8. SDS-PAGE (12,5%), com as fracções de purificação da AraL. Poço 1- Marcador de pesos moleculares, Poço 2- Extracto, Poço 3- Fração insolúvel (*pellet*), Poço 4- Fração solúvel (sobrenadante), Poço 5- Wash 1, Poço 6- Wash 2, Poço 7 a 9- Fração eluição 40mM imidazole. Poço 10- Fração de eluição 100 mM de imidazole. Mw(AraN)=29kDa.

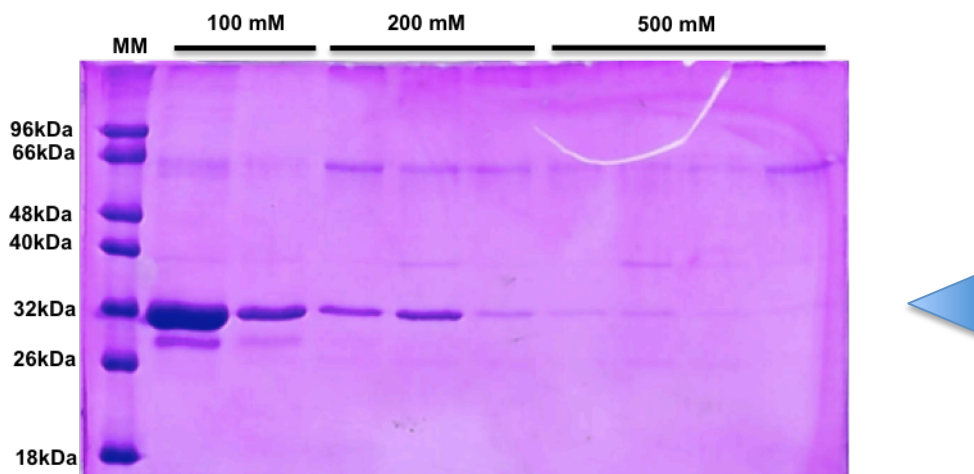


Figura III.9. SDS-PAGE (12,5%), com as fracções de purificação da AraL. Poço 1- Marcador de pesos moleculares, Poço 2 e 3- Fração 100 mM de imidazole, Poço 4 a 6- Fração de 200 mM de imidazole, Poço 7 a 9- Fracções 500 mM de imidazole. Mw(AraN)=29kDa.

Pela análise da figura III.8 é possível observar que há uma pequena quantidade de proteína que ainda assim fica insolúvel no entanto a grande maioria passa para a fração solúvel.

Confirmada a possibilidade de precipitar o SDS foi efectuado um novo ensaio, sendo que os passos de solubilização da AraL e precipitação do SDS foram acompanhados através de RMN e com recurso a SDS-PAGE com o intuito de perceber se haveria perdas de proteína em algum passo de purificação.

Numa primeira fase procedeu-se à precipitação do SDS e consequente recolha na forma de uma fração insolúvel (*pellet*). Na figura III.10 é possível confirmar que a fração insolúvel é composta na sua grande maioria por SDS.

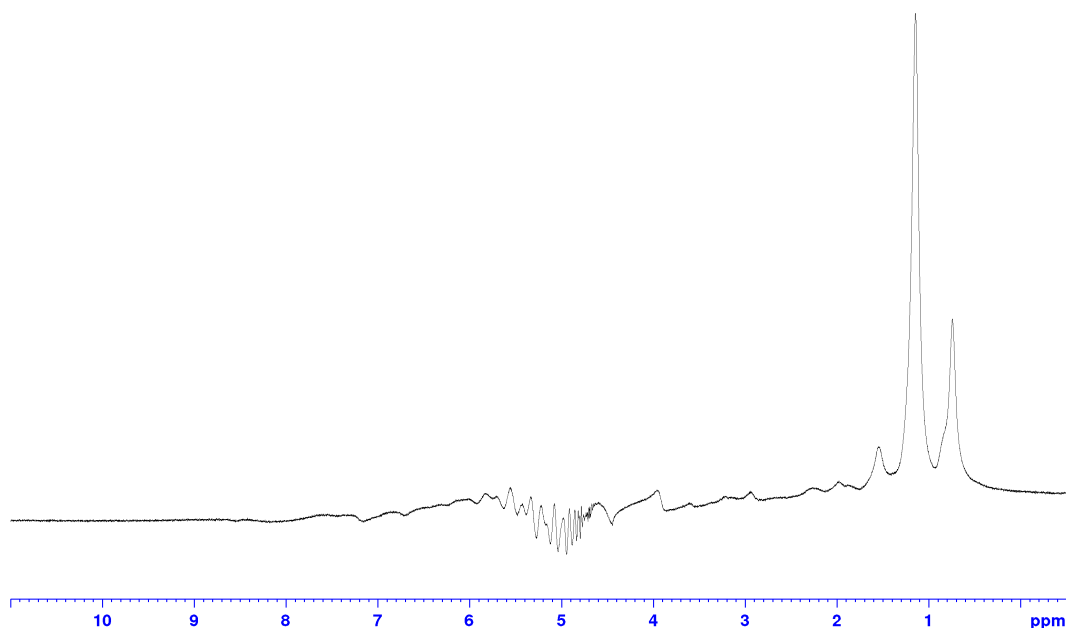


Figura III.10. Espectro ^1H da fração insolúvel. Após o SDS ser precipitado e subsequentemente sedimentado por centrifugação, é possível constatar a presença de ressonâncias espectrais associadas à cadeia alifática do dodecil sulfato.

Na figura III.11 está representado espectro ^1H da fração solúvel obtida. Neste é possível observar que eventualmente e devido ao alargamento das linhas presentes no espectro que algumas proteínas previamente solubilizadas com SDS recuperaram a sua estrutura tridimensional.

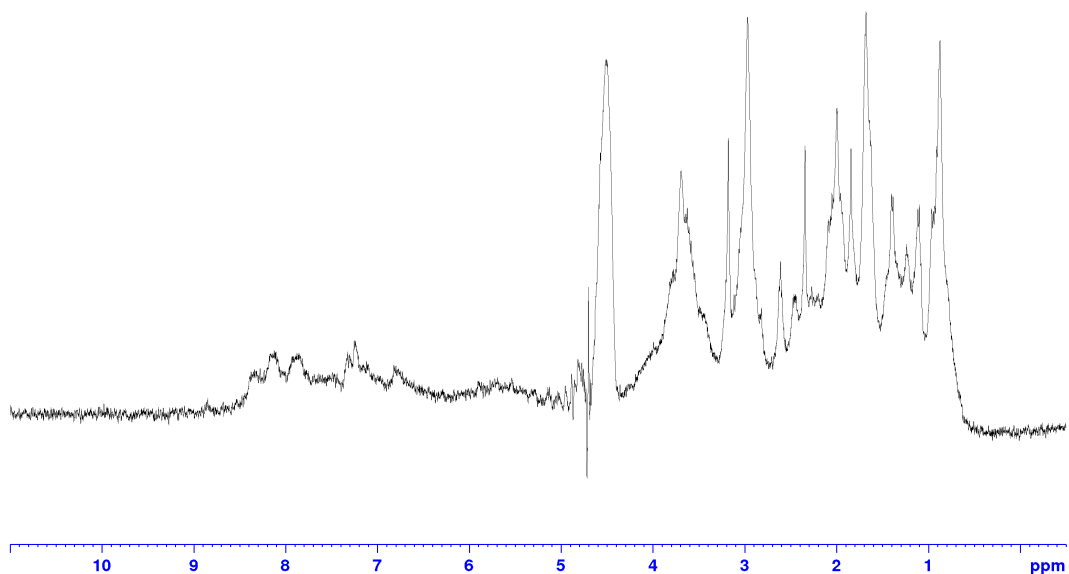


Figura III.11. Espectro ^1H da fração solúvel resultante do passo de precipitação do SDS.

Para perceber se a AraL tinha ficado retida na fração insolúvel juntamente com o SDS ou se a mesma se encontrava no sobrenadante foi efectuado um SDS-PAGE, ilustrado na figura III.12.

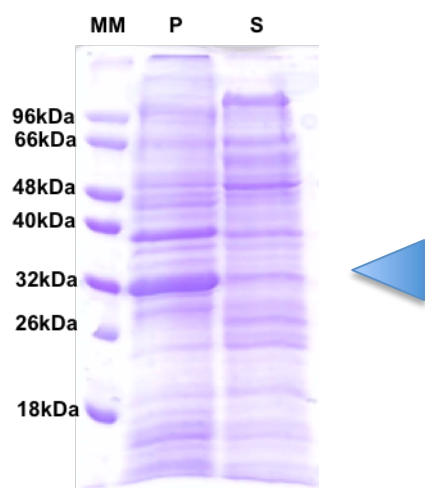


Figura II. 12. SDS-PAGE (12,5%), com as fracções de purificação da AraL. Poço 1- Marcador de pesos moleculares, Poço 2- Fração insolúvel, Poço 3- Fração solúvel. Mw(AraN)=29 kDa.

Através da análise da figura III.12, é possível constatar que a AraL ficou quase maioritariamente agarrada ao SDS na fração insolúvel, como é possível observar no poço 2 da mesma figura. No terceiro poço encontram-se um conjunto de outras proteínas que podem estar na origem do espectro da figura III.11, que à primeira vista parece apresentar proteínas na sua estrutura nativa ou tridimensional.

Com o objectivo de recuperar a proteína que se encontrava no *pellet* foram utilizadas duas metodologias: i) utilizar um agente surfactante que não o SDS para recuperar a AraL, sendo que neste caso foi utilizado 0,1% Sarkosyl, ii) subir o pH para 8,5 que anteriormente era de 7,4 para promover a desprotonação da AraL ($pI=6,0$) e consequentemente aumentar as repulsões electrostáticas entre o anião dodecil sulfato e a própria AraL. Na figura III.13 encontram-se os géis de SDS relativos a essa estratégia.

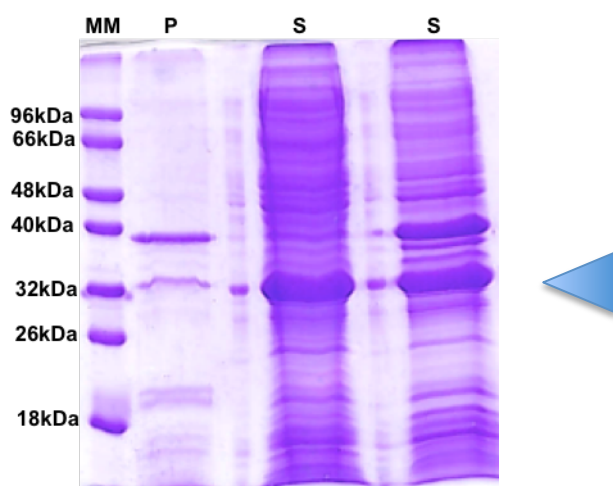


Figura III.13. SDS-PAGE (12,5%), com as fracções de purificação da AraL. Poço 1- Marcador de pesos moleculares, Poço 2- Fração insolúvel após tratamento do *pellet* com 0,1% sarkosyl a pH=7,4, Poço 3- Fração solúvel com dar AraL após tratamento com 0,1% sarkosyl a pH=7,4, Poço 4- Fração solúvel após solubilização do pellet obtido na Fração 3, utilizando 0,1% sarkosyl a pH=8,5.

Ainda no que à estratégia diz respeito, esta foi efectuada por duas etapas. Numa primeira fase a proteína que se encontrava no *pellet* juntamente com o SDS foi resuspendida com o tampão fosfato evidenciado na secção III.3.2.2.5, suplementado com 0,1% (p/v) sarkosyl. Esta mistura é subsequentemente centrifugada a 13 000 rpm's durante 15 minutos. A segunda estratégia faz uso do *pellet* anteriormente recolhido após tratamento com 0,1% (p/v) sarkosyl, sendo este resuspendido num tampão utilizado na primeira estratégia mas a pH=8,5. Seguidamente este é centrifugado a 13 000 rpm's durante 15 minutos, recolhendo-se o sobrenadante. Em ambas as estratégias, tal como é possível observar na figura III.13 a proteína ficou na fração solúvel.

De seguida a AraL foi purificada, sendo as fracções obtidas da colunas Ni-NTA encontram-se na figura III.14 e III.15.

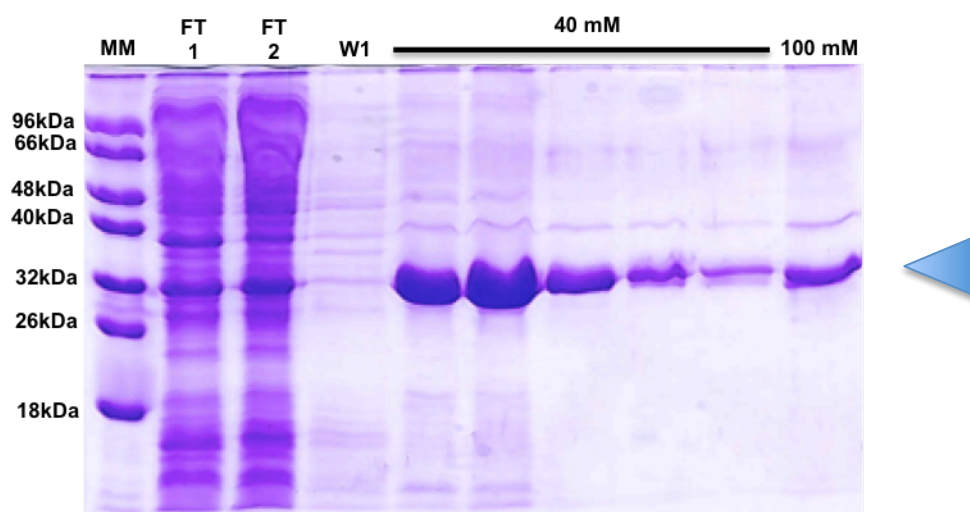


Figura III.14. SDS-PAGE (12,5%), com as fracções de purificação da AraL. Poço 1- Marcador de pesos moleculares, Poço 2- *Flow through* 1, Poço 3- *Flow through* 2, Poço 4- *Wash* 1, Poço 5 ao 9- Frações de eluição 40 mM de imidazole, Poço 10- Fração de eluição 100 mM de imidazole. Mw(AraL)= 29kDa.

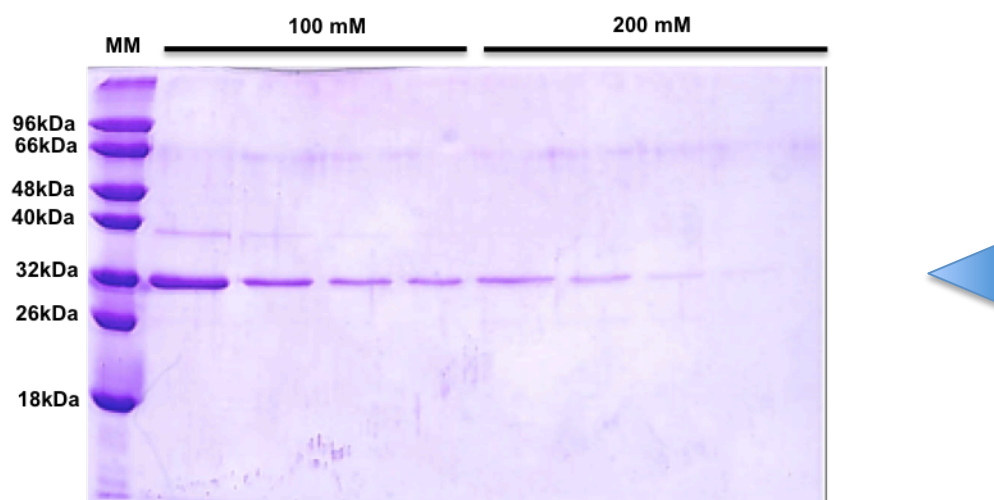


Figura III.15. SDS-PAGE (12,5%), com as fracções de purificação da AraL. Poço 1- Marcador de pesos moleculares, Poço 2 ao 5- Frações de eluição 100 mM de imidazole, Poço 6 ao 10- Frações de eluição 200 mM de imidazole. Mw(AraL)= 29 kDa.

Após purificação a atividade foi aferida como indicado na secção III.3.2.3, com as absorvâncias indicadas na tabela III.1.

Tabela III.1. Medição de atividade fosfatase da AraL. Medição efectuada após tratamento com SDS e subsequente renaturação.

Diluições	Absorvância
1:1	0,211
	0,214
	0,196
1:2	0,131
	0,127
	0,111
1:4	0,086
	0,091
	0,084

III.3.3.1.4. Solubilização e renaturação da AraL com recurso a ureia

A AraL na forma de corpos de inclusão é solubilizada com recurso a ureia (secção III.3.2.2.3). O processo de renaturação é efectuada através da redução da concentração de ureia (6M para 0M). O gel apenas com as fracções de purificação está representado na figura III.16. A AraL é eluída a 200 mM de imidazole (poço 13).

Após renaturação foi efectuada um teste qualitativo para verificar se a AraL recuperou a sua estrutura nativa, e consequentemente atividade enzimática. Para tal foi utilizado o protocolo da secção III.3.2.3, tendo sido efectuada uma única alteração ao protocolo, com a AraL a ser incubada a 37°C. O processo de solubilização funcionou, visto que a AraL ficou solúvel na fracção de eluição 13, tendo potencialmente evidenciado atividade fosfatase com Abs₄₀₅ de 0,02.

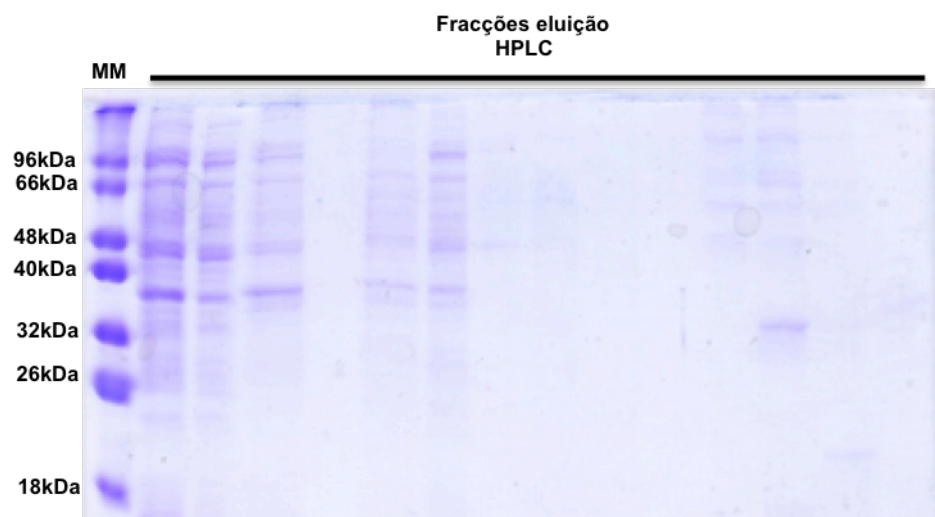


Figura III.16. SDS-PAGE (12,5%), com as fracções de purificação da AraL. Poço 1- Marcador de pesos moleculares, Poço 2 a 10-Fracções de eluição, gradiente contínuo de imidazole. Mw(AraL)= 29kDa.

III.3.3.1.5. Solubilização e renaturação da AraL com recurso a GmdCI

Tal como na secção anterior, foi também efectuada solubilização da AraL na forma de corpos de inclusão com recurso a GmdCI. Como é possível constatar na figura, a AraL é solubilizada e eluída nas fracções 10 a 13 no gel representado na figura III.17.

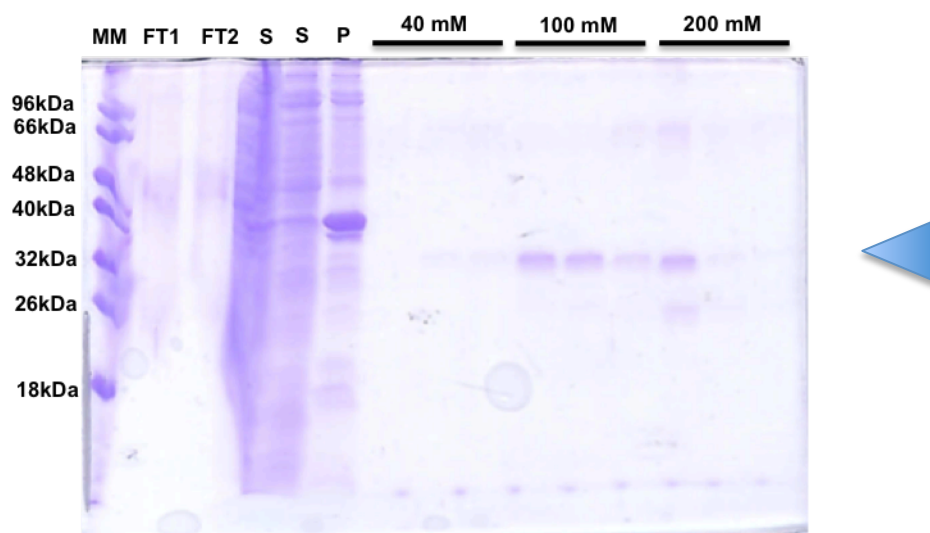


Figura III.17. SDS-PAGE (12,5%), com as fracções de purificação da AraL. Poço 1- Marcador de pesos moleculares, Poço 2 e 3- *Flow Through*, Poço 4 e 5- Fração solúvel, Poço 6- Fração insolúvel, Poço 7 a 9- Fração de eluição 40 mM imidazole, Poço 10 a 12- Fração de eluição 100 mM imidazole, Poço 13 a 15- Fração de eluição 200 mM de imidazole. Mw(AraL)= 29kDa.

Tal como na secção anterior foram medidas as absorvâncias da AraL renaturada na presença de pNPP (tabela III.2). Tendo-se concluído, que tal como anteriormente, tanto na solubilização dos copos de inclusão com SDS bem como ureia, esta recupera a sua atividade.

Tabela III.2. Medição de atividade fosfatase da AraL. Medição efectuada após tratamento com GuHCL e subsequente renaturação.

Diluições	Absorvância
1:1	0,020
	0,022
	0,025
1:2	0,014
	0,013
	0,011
1:4	0,08
	0,08
	0,09

III.4. Conclusões

Ao longo deste capítulo foram desenvolvidos diferentes protocolos com o objectivo de solubilizar a AraL, que à partida se encontrava na sua forma insolúvel sobre a forma de corpos de inclusão. Para tal desenvolveram-se metodologias de solubilização com recurso a detergentes tais como o SDS e agentes caotrópicos tais como ureia e GmdCl. Após solubilizada a AraL é renaturada através da precipitação do SDS e diálisada com o objectivo de remover excipientes de SDS, ou através de remoção de ureia ou GmdCl numa coluna IMAC NI-NTA.

Através dos resultados obtidos, é possível concluir que o SDS pode ser removido pela co-precipitação do mesmo com auxílio a sais de potássio tais como o KH_2PO_4 à concentração de 150 mM, processo este auxiliado pela incubação da mistura reacional em gelo. Os espectros de ^1H são indicativos que após a remoção da SDS, ocorre recuperação da estrutura nativa das proteínas em questão, dados que podem ser apoiados pelo facto da AraL ter recuperado a sua atividade para com o substrato pNPP.

Por outro lado, o uso de agentes caotrópicos tais como a ureia e o GmdCl, revelaram-se eficientes no processo de solubilização da AraL na forma de corpos de inclusão, no entanto, visto que os crescimentos celulares não foram efectuados nas mesmas condições (volume reacional) e na medição das atividades não foi tido em conta a concentração de enzima (AraL) é impossível de um modo quantitativo avaliar quais dos processos é mais rentável quer de um ponto de vista de rendimento de proteína solúvel obtida quer de fração de proteína renaturada.

IV. Conclusões gerais e perspectivas futuras

Os principais objectivos desta tese residiram na contribuição para uma melhor compreensão dos mecanismos de reconhecimento molecular associados a duas proteínas codificadas no operão de metabolização de arabinose presente na *Bacillus subtilis*.

Relativamente à AraN, através de estudos de STD-RMN foi possível aferir que a subunidade procedente à subunidade redutora em arabinoligossacáridos apresenta um papel fulcral no reconhecimento dos mesmos. Por um lado, é esta a subunidade que apresenta um maior grau de saturação ao longo do estudo da série de interações AraN-arabinoligossacáridos (de duas a quatro subunidades). Por outro lado, o facto da arabinose não apresentar STD contrariamente ao observado para a arabinobiose, pode ser indicativo de que a adição de uma segunda subunidade é fundamental no estabelecimento de interações com a AraN. Através de um modelo de homologia da AraN, foi possível efetuar *dockings* moleculares para a arabinotriose. A análise destes modelos permitiu por uma lado tirar conclusões acerca dos resíduos da proteína que se encontram a interagir com a arabinotriose, e por outro lado aferir que regiões da arabinotriose se encontram hipoteticamente em maior proximidade com a AraN. O processo de reconhecimento, tal como previamente observado noutras SBP, é governado por um conjunto de resíduos que irão estabelecer interações com grupos – CH do arabinoligossacárido, através do *stacking* de cadeias laterais aromáticas nos anteriores. Na AraN estes resíduos são o Y54, W253 e Y254, sendo que destes o resíduo W253 é conservado nesta família de proteínas encontrando-se a envolvido no processo de reconhecimento molecular dos substratos em questão. As cadeias laterais destes três aminoácidos parecem estar a interagir com as três subunidades que compõem a arabinotriose, de um modo singular na medida que cada cadeia lateral parece apenas interagir com um monómero (1 para 1). O facto da maior amplificação de STD ser conseguida com a arabinotriose, pode ser indicativo de que o local de ligação está optimizado para o reconhecimento de oligosacáridos compostos por três subunidades, algo que vai ao encontro do facto do local de ligação se encontrar decorado pelo três resíduos aromáticos acima evidenciados. Para além das interações em cima mencionadas, um dos modelos de *docking* também previa a formação de pontes de hidrogénio entre as subunidades redutoras e não redutoras com a AraN através dos resíduos E175, D178 e E78 respectivamente.

Uma das principais inconsistências dos estudos efectuados por STD-RMN está associada a fenómenos de relaxação longitudinal T_1 . Tal como é ilustrado na literatura [98], estes carboidratos apresentam algumas variações de T_1 associadas aos diferentes prótons que o compõem, variações estas que podem ter implicações na determinação inequívoca de um modo de ligação através do mapeamento de epítomos, pois num determinado tempo de saturação, os efeitos de relaxação não irão ser iguais para os diferentes prótons podendo ocorrer uma subvalorização ou sobrevalorização da interação caso o seu T_1 seja inferior ou superior ao tempo de saturação respectivamente. A determinação de um parâmetro STD_{fit} , este último determinado em condições nas quais os efeitos de T_1 são negligenciáveis, originou um mapeamento diferente na medida que o próton H4 α da subunidade redutora apresenta maior grau de saturação. Este dado poderá fazer sentido, visto que no modelo de *docking*, o próton que está em maior proximidade com a AraN é o H4 α . A melhor

maneira de aferir o impacto que as relaxações longitudinais podem ter no mapeamento de epítomos, é através do cálculo do T_1 de cada próton através de uma experiência de *inversion recovery*.

Com recurso à técnica de STD-RMN, foi ainda possível constatar que afinidade da AraN não é tão grande quanto o reportado para outras SBP's, visto que a primeira está na ordem dos 200 a 700 μM , ao passo que as segundas apresentam um K_d na ordem de 1 a 10 μM . Por outro lado a AraN não é apenas específica para arabinoligossacáridos. Em condições de competição a arabinotriose parece competir com a celotetraose para o mesmo local de ligação. O facto de se verificar esta competição ao nível da AraN que se encontra envolvida num transportador dedicado ao importe de arabinoligossacáridos, pode ser uma forma de repressão catabólica exercida directamente ao nível do transportador. Em *Bacillus subtilis*, é conhecida a existência de um transportador dedicado ao transporte de celobiose codificado ao nível do operão *lic*. De um ponto de vista biológico e metabólico a incorporação de celoligossacáridos pode ser energeticamente mais favorável, assim em condições de competição os celoligossacáridos constituídos por unidades monoméricas de glucose irão funcionar como fonte de carbono primária. Uma maneira de verificar esta hipótese seria através da medição do transporte destas fontes de carbono marcadas isotopicamente em condições de competição, nas quais um dos carboidratos é marcado e o segundo não, e subsequente determinação da sua localização sub-celular. Por outro lado, o facto da maltohexaose não competir com a arabinotriose pode estar associado a mecanismos estereoespecíficos ou face ao seu tamanho. Para se perceber melhor os motivos pelos quais a celotetraose compete com a arabinotriose ao contrário da maltohexaose é necessário efetuar experiências de competição com oligossacáridos com igual tamanho (arabinotriose, celotriose, maltotriose).

Os estudos *in silico*, nomeadamente através de *docking* molecular e alinhamentos, serviram como foi em cima indicado, para estabelecer resíduos putativos de se encontrarem a interagir ao nível do local de ligação da AraN, e em geral para estabelecer um modo de reconhecimento arabinotriose-AraN. Face à abertura da cavidade de ligação, foram observados essencialmente dois polos nos quais se encontram resíduos passíveis de estarem envolvidos no estabelecimento de interações com a arabinotriose. Para tal foram efectuados dois mutantes R81A e W253A, sendo que este último ainda não foi testado. Através de STD-RMN, não se verificou alterações no que toca ao padrão de reconhecimento da arabinotriose quando comparado com o padrão efectuado com a AraN(wt). Assim, é essencial analisar por um lado quais as implicações de um estudo semelhante ao nível da AraN(W253A) e da AraN(W253A e R81A), para perceber se a mutação no primeiro leva à abolição completa do reconhecimento da arabinotriose, ou no segundo se o processo de reconhecimento é cooperativo, isto é, se o processo de reconhecimento envolve ambos os resíduos que por um lado se encontram a estabelecer interações de diferentes naturezas, hidrofóbicas no caso do W253 e polares no caso do R81, e por outro lado visto que se encontram localizados em lobos diferentes da AraN. Ainda relativamente aos estudos *in silico*, estes foram efectuados sem que o receptor molecular tivesse mobilidade, e tendo em conta que associado ao processo de reconhecimento molecular está a deslocação do domínio N-terminal e fecho sobre o local de ligação, é também essencial mimetizar esta condição através da modelação de um *docking* com o receptor molecular móvel.

Para além da proteína AraN (SBP), um dos objectivos da tese era o de estudar os mecanismos de reconhecimento molecular nível sobre o ponto de vista da fosfatase AraL. Visto que esta proteína era sobreexpressa inicialmente sobre a forma de corpos de inclusão foi desenvolvida uma metodologia com recurso a SDS para solubilizar esta proteína, sendo que o detergente é posteriormente precipitado com 150mM de KH_2PO_4 e a proteína extraída com 0,1% de Sarkosyl. Futuramente a renaturação desta proteína pode ser aferida com recurso à aquisição de um espectro de HSQC, e a sua atividade com recurso a uma titulação com um substrato fosforilado.

V. Bibliografia

1. Dwek, R.A., *Glycobiology: Toward Understanding the Function of Sugars*. Chem Rev, 1996. **96**: p. 683-720.
2. Gabius, H.J., et al., *Chemical biology of the sugar code*. Chembiochem, 2004. **5**(6): p. 740-64.
3. Rosen, S.D., Bertozzi, C. R., *The selectins and their ligands*. Curr Opin Cell Biol, 1994. **6**(5): p. 663-73.
4. Sauter, N.K., Bednarski, M. D., Wurzburg, B. A., Hanson, J. E., Whitesides, G. M., Skehe, J. J., Wiley, D. C., *Binding of influenza virus hemagglutinin to analogs of its cell-surface receptor, sialic acid: analysis by proton nuclear magnetic resonance spectroscopy and x-ray crystallography*. Biochemistry, 1992. **31**(40): p. 9609-9621.
5. Borén, T., Falk, P., Roth, K. A., Larson, G., Normark, S., *Attachment of Helicobacter pylori to human gastric epithelium mediated by blood group antigens*. Science, 1993. **262**(5141): p. 1892-5.
6. Fantini, J., Cook, D. G., Nathanson, N., Spitalnik, S. L., Gonzalez- Scarano, F., *Infection of colonic epithelial cell lines by type 1 human immunodeficiency virus is associated with cell surface expression of galactosylceramide ,a potential alternative gpl20 receptor*. Proc Natl Acad Sci U S A, 1993. **90**: p. 2700-2704.
7. Rother, R.P., Squinto, S. P., *The α -Galactosyl Epitope: A Sugar Coating That Makes Viruses and Cells Unpalatable*. Cell, 1996. **86**(2): p. 185-188.
8. Holgersson, J., A. Gustafsson, and M.E. Breimer, *Characteristics of protein-carbohydrate interactions as a basis for developing novel carbohydrate-based antirejection therapies*. Immunol Cell Biol, 2005. **83**(6): p. 694-708.
9. Ferreira, M.J. and I. Sa-Nogueira, *A multitask ATPase serving different ABC-type sugar importers in Bacillus subtilis*. J Bacteriol, 2010. **192**(20): p. 5312-8.
10. Godinho, L.M., Sá-Nogueira, I., *Characterization and regulation of a bacterial sugar phosphatase of the haloalkanoate dehalogenase superfamily, AraL, from Bacillus subtilis*. FEBS J, 2011. **278**: p. 2511-2524.
11. Stulke, J., Hillen, W., *Regulation of carbon catabolism in Bacillus species*. Annu Rev Biochem, 2000. **54**: p. 849-80.
12. Carpita, N.C., Gibeaut, D. M., *Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth*. Plant J., 1993. **3**(1): p. 1-30.
13. Beldman, G.S., H. A., Pitson, S. M., Searle-van Leeuwen, M . J. F., Voragen, A. G. J., *Arabinans and arabinan degrading enzymes*. Advances in Macromolecular Carbohydrate Research, 1997. **1**: p. 1-64.
14. Leal, T.F. and I. de Sa-Nogueira, *Purification, characterization and functional analysis of an endo-arabinanase (AbnA) from Bacillus subtilis*. FEMS Microbiol Lett, 2004. **241**(1): p. 41-8.
15. Inacio, J.M. and I. de Sa-Nogueira, *Characterization of abn2 (yxiA), encoding a Bacillus subtilis GH43 arabinanase, Abn2, and its role in arabino-polysaccharide degradation*. J Bacteriol, 2008. **190**(12): p. 4272-80.

16. Inacio, J.M., I.L. Correia, and I. de Sa-Nogueira, *Two distinct arabinofuranosidases contribute to arabino-oligosaccharide degradation in Bacillus subtilis*. Microbiology, 2008. **154**(Pt 9): p. 2719-29.
17. Sá-Nogueira, I., Lencastre, H., *Cloning and characterization of araA, araB and araD, the structural genes for L-arabinose utilization in Bacillus subtilis*. J. Bacteriol, 1989. **171**: p. 4088-91.
18. Mueckler, M., Caruso, C., Baldwin, S. A., Panico, M., Blench, I., Morris, H. R., Allard, W. J., Lienhard, G. E., Lodish, H. F., *Sequence and structure of a human glucose transporter*. Science, 1985. **229**(4717): p. 941-945.
19. Schonert, S., et al., *Maltose and maltodextrin utilization by Bacillus subtilis*. J Bacteriol, 2006. **188**(11): p. 3911-22.
20. Sorensen, T.L., J.V. Moller, and P. Nissen, *Phosphoryl transfer and calcium ion occlusion in the calcium pump*. Science, 2004. **304**(5677): p. 1672-5.
21. Abramson, J., et al., *Structure and mechanism of the lactose permease of Escherichia coli*. Science, 2003. **301**(5633): p. 610-5.
22. Goswitz, V.C., Brooker, R. J., *Structural features of the uniporter/symporter/antiporter superfamily*. Protein Sci, 1995. **4**: p. 534-537.
23. Lingrel, J.B., Kuntzweiler, T., *Na⁺, K⁺-ATPase*. J Biol Chem, 1994. **269**(31): p. 19659-62.
24. International Transporter, C., et al., *Membrane transporters in drug development*. Nat Rev Drug Discov, 2010. **9**(3): p. 215-36.
25. Saier, M.H., Jr., *Families of proteins forming transmembrane channels*. J Membr Biol, 2000. **175**(3): p. 165-80.
26. Saier, M.H., Jr. and I.T. Paulsen, *Phylogeny of multidrug transporters*. Semin Cell Dev Biol, 2001. **12**(3): p. 205-13.
27. Davidson, A.L. and P.C. Maloney, *ABC transporters: how small machines do a big job*. Trends Microbiol, 2007. **15**(10): p. 448-55.
28. Oldham, M.L., A.L. Davidson, and J. Chen, *Structural insights into ABC transporter mechanism*. Curr Opin Struct Biol, 2008. **18**(6): p. 726-33.
29. Higgins, C.F., *ABC transporters: from micro to man*. Annu Rev Cell Biol, 1992. **8**: p. 67-113.
30. Moitra, K., H. Lou, and M. Dean, *Multidrug efflux pumps and cancer stem cells: insights into multidrug resistance and therapeutic development*. Clin Pharmacol Ther, 2011. **89**(4): p. 491-502.
31. Dawson, R.J. and K.P. Locher, *Structure of a bacterial multidrug ABC transporter*. Nature, 2006. **443**(7108): p. 180-5.
32. Fletcher, J.I., et al., *ABC transporters in cancer: more than just drug efflux pumps*. Nat Rev Cancer, 2010. **10**(2): p. 147-56.
33. Rees, D.C., E. Johnson, and O. Lewinson, *ABC transporters: the power to change*. Nat Rev Mol Cell Biol, 2009. **10**(3): p. 218-27.
34. Davidson, A.L., et al., *Structure, function, and evolution of bacterial ATP-binding cassette systems*. Microbiol Mol Biol Rev, 2008. **72**(2): p. 317-64, table of contents.

35. Locher, K.P., *Review. Structure and mechanism of ATP-binding cassette transporters*. Philos Trans R Soc Lond B Biol Sci, 2009. **364**(1514): p. 239-45.
36. Borths, E.L., et al., *The structure of Escherichia coli BtuF and binding to its cognate ATP binding cassette transporter*. Proc Natl Acad Sci U S A, 2002. **99**(26): p. 16642-7.
37. Karpowich, N.K., et al., *Crystal structures of the BtuF periplasmic-binding protein for vitamin B12 suggest a functionally important reduction in protein mobility upon ligand binding*. J Biol Chem, 2003. **278**(10): p. 8429-34.
38. Shuman, H.A., *Active transport of maltose in Escherichia coli K12. Role of the periplasmic maltose-binding protein and evidence for a substrate recognition site in the cytoplasmic membrane*. Journal of Biological Chemistry, 1982. **257**: p. 5455-5461.
39. Treptow, N.A., Shuman, H. A., *Genetic evidence for substrate and periplasmic-binding-protein recognition by the MalF and MalG proteins, cytoplasmic membrane components of the Escherichia coli maltose transport system*. J. Bacteriol, 1985. **163**(2): p. 654-60.
40. Neu, H.C., Heppel, L. A., *The Release of Enzymes from Escherichia coli by Osmotic Shock and during the Formation of Spheroplasts*. J. Biol. Chem, 1965. **240**: p. 3685-3692.
41. Gilson, E., Alloing, G., Schmidt, T., Claverys, J-P., Dudler, R., Hofnung, M., *Evidence for high affinity binding-protein dependent transport systems in Gram-positive bacteria and in Mycoplasma*. EMBO Journal, 1988. **7**(12): p. 3971-3974.
42. Sutcliffe, I.C., Russell, R. R., *Lipoproteins of gram-positive bacteria*. J. Bacteriol, 1995. **177**(5): p. 1123-1128.
43. Tiemen van der Heide, P., B., *ABC transporters: one, two or four extracytoplasmic substrate-binding sites?* EMBO Reports, 2002. **3**(10): p. 938-43.
44. Biemans-Oldehinkel, E., Poolman, B., *On the role of the two extracytoplasmic substrate-binding domains in the ABC transporter OpuA*. EMBO Journal, 2003. **22**(22): p. 5983-5993.
45. Cuneo, M.J., et al., *The backbone structure of the thermophilic Thermoanaerobacter tengcongensis ribose binding protein is essentially identical to its mesophilic E. coli homolog*. BMC Struct Biol, 2008. **8**: p. 20.
46. Hsiao, C.D., Sun, Y. J., Rose, J. & Wang, B. C., *The Crystal Structure of Glutamine-binding Protein from Escherichia coli*. J. Biol. Chem, 1996. **278**: p. 8429-8434.
47. Levdivkov, V.M., et al., *The structure of the oligopeptide-binding protein, AppA, from Bacillus subtilis in complex with a nonapeptide*. J Mol Biol, 2005. **345**(4): p. 879-92.
48. Sleight, S.H., Tame, J. R. H., Dodson, E. J., Wilkinson, A. J., *Peptide Binding in OppA, the Crystal Structures of the Periplasmic Oligopeptide Binding Protein in the Unliganded Form and in Complex with Lysyllysine*. Biochemistry, 1997. **36**: p. 9747-9758.
49. Yao, N., Ledvina, P. S., Choudhary, A., Quirocho, F. A., *Modulation of a Salt Link Does Not Affect Binding of Phosphate to Its Specific Active Transport Receptor*. Biochemistry, 1996. **35**: p. 2079-2085.
50. Bruns, C.M., Nowalk, A. J., Arvai, A. S., McTigue, M. A., Vaughan, K. G., Mietzner, T. A., McRee, D. E., *Structure of Haemophilus influenzae Fe(+3)-binding protein reveals convergent evolution within a superfamily*. Nat Struct Biol, 1997. **4**(11): p. 919-24.

51. Silva, Z., et al., *The high-affinity maltose/trehalose ABC transporter in the extremely thermophilic bacterium Thermus thermophilus HB27 also recognizes sucrose and palatinose.* J Bacteriol, 2005. **187**(4): p. 1210-8.
52. Alloing, G., de Philip, P., Claverys, J.-P., *Three highly homologous membrane-bound lipoproteins participate in oligopeptide transport by the Ami system of Streptococcus pneumoniae.* J Mol Biol, 1994. **241**: p. 44-58.
53. Quioco, F.A. and C.F. Higgins, *Atomic Structures of Periplasmic Binding Proteins and the High-Affinity Active Transport Systems in Bacteria [and Discussion].* Philosophical Transactions of the Royal Society B: Biological Sciences, 1990. **326**(1236): p. 341-352.
54. Abbott, D.W., et al., *The molecular basis of glycogen breakdown and transport in Streptococcus pneumoniae.* Mol Microbiol, 2010. **77**(1): p. 183-99.
55. Quioco, F.A., Ledvina, P. S., *Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis variation of common themes.* Mol Microbiol, 1996. **20**(1): p. 17-25.
56. Fukami-Kobayashi, K., Tateno, Y., Nishikawa, K, *Domain dislocation- a change of core structure in periplasmic binding proteins in their evolutionary.* J Mol Biol, 1999. **286**: p. 279-290.
57. Mao, B., Pear, M. R., McCammon, J. A., Quioco, F. A., *Hinge-bending in L-arabinose-binding protein. The "Venus's-flytrap" model.* J Biol Chem, 1982. **257**: p. 1131-1133.
58. Bjorkman, A.J., Mowbray, S. L., *Multiple open forms of ribose-binding protein trace the path of its conformational change.* J Mol Biol, 1998. **279**: p. 651-664.
59. Magnusson, U., et al., *Hinge-bending motion of D-allose-binding protein from Escherichia coli: three open conformations.* J Biol Chem, 2002. **277**(16): p. 14077-84.
60. Quioco, F.A., *Carbohydrate-binding proteins: Tertiary structures and protein-sugar interactions.* Annu Rev Biochem, 1986. **55**: p. 287-315.
61. Quioco, F.A., *Protein -carbohydrate interactions: basic molecular features.* Pure and Applied Chemistry, 1989. **61**: p. 1293-1306.
62. Yao, N., Trakhanov, S., Quioco, F. A., *Refined 1.89-A Structure of the Histidine-Binding Protein Complexed with Histidine and Its Relationship with Many Other Active Transport/Chemosensory Proteins.* Biochemistry, 1994. **33**: p. 4769-4779.
63. Oh, B.-H., Pandit, J., Kang, C-Hee., Nikaido, K., Gocken, S., Ames, G. F-L., Kim, S-Hou., *Three-dimensional Structures of the Periplasmic Lysine/Arginine/Ornithine-binding Protein with and without a Ligand.* JBC, 1993. **268**(15): p. 11348-11355.
64. Johnson, L.N., *The regulation of protein phosphorylation.* Biochem Soc Trans, 2009. **37**(Pt 4): p. 627-41.
65. Allen, K.N. and D. Dunaway-Mariano, *Phosphoryl group transfer: evolution of a catalytic scaffold.* Trends Biochem Sci, 2004. **29**(9): p. 495-503.
66. Lu, Z., D. Dunaway-Mariano, and K.N. Allen, *The catalytic scaffold of the haloalkanoic acid dehalogenase enzyme superfamily acts as a mold for the trigonal bipyramidal transition state.* Proc Natl Acad Sci U S A, 2008. **105**(15): p. 5687-92.

67. Meyer, B., Peters, T., *NMR Spectroscopy Techniques for Screening and Identifying Ligand Binding to Protein Receptors*. Angew. Chem. Int, 2003. **42**: p. 864-890.
68. Shuker, S.B., Hajduk, P. J., Meadows, R. P., Fesik, S. W., *Discovery high-affinity ligands for proteins: SAR by NMR*. Science, 1996. **274**(5292): p. 1531-34.
69. Medek, A., Hajduk, P. J., Mack, J., Fesik, S. W., *The Use of Differential Chemical Shifts for Determining the Binding Site Location and Orientation of Protein-Bound Ligands*. J Am Chem Soc, 2000. **122**: p. 1241-42.
70. Gardner, K.H., Zhang, X., Gehring, K., Kay, L. E., *Solution NMR Studies of a 42 KDa Escherichia Coli Maltose Binding Protein/B-Cyclodextrin Complex: Chemical Shift Assignments and Analysis*. J Am Chem Soc, 1998. **120**: p. 11738-11748.
71. Tugarinov, V., Muhandiram, R., Ayed, A., Kay, L. E., *Four-Dimensional NMR Spectroscopy of a 723-Residue Protein: Chemical Shift Assignments and Secondary Structure of Malate Synthase G*. J Am Chem Soc, 2002. **124**: p. 10025-35.
72. Hajduk, P.J., Gerfin, T., Boehlen, J-M., Haberli, M., Marek, D., Fesick, S. W., *High-Throughput Nuclear Magnetic Resonance-Based Screening*. J Med Chem, 1999. **42**: p. 2315-17.
73. Frutos, S., et al., *Disruption of the HIV-1 protease dimer with interface peptides: structural studies using NMR spectroscopy combined with [2-(13)C]-Trp selective labeling*. Biopolymers, 2007. **88**(2): p. 164-73.
74. Medek, A., Olejniczak, E. T., Meadows, R. P., Fesik, S. W. , *An approach for high-throughput structure determination of proteins by NMR spectroscopy*. Journal of Biomolecular NMR, 2000. **18**: p. 229-238.
75. Pellecchia, M., Meininger, D., Dong, Q., Chang, E., Jack, R., Sem, D. S., *NMR-based structural characterization of large protein-ligand interactions*. Journal of Biomolecular NMR, 2002. **22**: p. 165-73.
76. Salvatella, X., et al., *A tetraguanidinium ligand binds to the surface of the tetramerization domain of protein P53*. Angew Chem Int Ed Engl, 2004. **43**(2): p. 196-8.
77. Pellecchia, M., *Solution nuclear magnetic resonance spectroscopy techniques for probing intermolecular interactions*. Chem Biol, 2005. **12**(9): p. 961-71.
78. Tarrago, T., et al., *A cost-effective labeling strategy for the NMR study of large proteins: selective 15N-labeling of the tryptophan side chains of prolyl oligopeptidase*. Chembiochem, 2009. **10**(17): p. 2736-9.
79. Weigelt, J., Dongen, M., Uppenberg, J., Schultz, J., Wikstrom, M., *Site-Selective Screening by NMR Spectroscopy with Labeled Amino Acid Pairs*. J Am Chem Soc, 2001. **124**(11): p. 2446-47.
80. Pellecchia, M., Meininger, D., Shen, A. L., Jack, R., Kasper, B. C., Sem, D. S., *SEA-TROSY (Solvent Exposed Amides with TROSY): A Method to Resolve the Problem of Spectral Overlap in Very Large Proteins*. J Am Chem Soc, 2001. **123**: p. 4633-34.
81. Jahnke, W., S. Rüdisser, and M. Zurini, *Spin Label Enhanced NMR Screening*. Journal of the American Chemical Society, 2001. **123**(13): p. 3149-3150.

82. Tang, C., C.D. Schwieters, and G.M. Clore, *Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR*. Nature, 2007. **449**(7165): p. 1078-82.
83. Sauter, N.K., Bednarski, M. D., Wurzburg, B. A., Hanson, J. E., Whitesides, G. M., Skehe, J. J., Wiley, D. C., *Hemagglutinins from Two Influenza Virus Variants Bind to Sialic Acid Derivatives with Millimolar Dissociation Constants: A 500-MHz Proton Nuclear Magnetic Resonance Study*. Biochemistry, 1989. **28**: p. 8388-96.
84. Schmitt, T.H., Zheng, Z., Jardetzky, O., *Dynamics of Tryptophan Binding to Escherichia coli Trp Repressor Wild Type and AV77 Mutant: An NMR Study*. Biochemistry, 1995. **34**: p. 13183-89.
85. Roldos, V., F.J. Canada, and J. Jimenez-Barbero, *Carbohydrate-protein interactions: a 3D view by NMR*. Chembiochem, 2011. **12**(7): p. 990-1005.
86. Ni, F., *Recent developments in transferred NOE methods*. Prog Nucl Magn Reson Spectrosc, 1994. **26**: p. 517-606.
87. Hajduk, P.J., Olejniczak, E. T., Fesick, S. W., *One-Dimensional Relaxation- and Diffusion-Edited NMR Methods for Screening Compounds That Bind to Macromolecules*. J Am Chem Soc, 1997. **119**: p. 12257-61.
88. Mayer, M., Meyer, B., *Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy*. Angew. Chem. Int, 1999. **38**(12): p. 1784-1788.
89. Meyer, B., Klein, J., Mayer, M., Meinecke, R., Moller, H., Nefte, A., Schuster, O., Wulfken, J., Ding, Y., Knaie, O., Labbe, J., Palcic, M. M., Hindsgaul, O., Wagner, B., Ernst, B., *Saturation transfer difference NMR spectroscopy for identifying ligand epitopes and binding specificities*. Ernst Schering Res. Found. Workshop, 2004: p. 149–167
90. Dalvit, C., Pevarello, P., Tatò, M., Veronesi, M., Vulpetti, A., Sundstrom, M., *Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water*. Journal of Biomolecular NMR, 2000. **18**: p. 65-68.
91. Dalvit, C., Fogliatto, G., Stewart, A., Veronesi, M., Stockman, B., *WaterLOGSY as a method for primary NMR screening: Practical aspects and range of applicability*. Journal of Biomolecular NMR, 2001. **21**: p. 349-359.
92. Price, W.S., *Pulsed-Field Gradient Nuclear Magnetic Resonance as a Tool for Studying Translational Diffusion: Part 1. Basic Theory*. Concepts Mag Reson, 1997. **9**: p. 299-336.
93. Price, W.S., *Pulsed-Field Gradient Nuclear Magnetic Resonance as a Tool for Studying Translational Diffusion: Part II. Experimental Aspects*. Concepts Mag Reson, 1998. **10**(4): p. 197-237.
94. Dehner, A. and H. Kessler, *Diffusion NMR spectroscopy: folding and aggregation of domains in p53*. Chembiochem, 2005. **6**(9): p. 1550-65.
95. Viegas, A., et al., *Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding*. Journal of Chemical Education, 2011. **88**(7): p. 990-994.
96. Studier, F.W., *Protein production by auto-induction in high-density shaking cultures*. Protein Expression and Purification, 2005. **41**(1): p. 207-234.

97. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding*. Anal Biochem, 1976. **72**: p. 248-54.
98. Cros, S., Imbert, A., Bouchemal, N., Du Penhoat, C. H., Perez, S., *Modeling of arabinofuranose and arabinan, II. Nmr and Conformational analysis of arabinobiose and arabinan*. Biopolymers, 1994. **34**: p. 1433-1447.
99. Viegas, A.J.M., *Molecular determinants of ligand specificity in carbohydrate-binding modules: an NMR and X-ray crystallography integrated study*, in Chemistry Department. 2012, FCT-UNL.
100. Zhang, Y., *Template-based modeling and free modeling by I-TASSER in CASP7*. Proteins, 2007. **69 Suppl 8**: p. 108-17.
101. Roy, A., A. Kucukural, and Y. Zhang, *I-TASSER: a unified platform for automated protein structure and function prediction*. Nat Protoc, 2010. **5**(4): p. 725-38.
102. Morris, G.M., et al., *AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility*. J Comput Chem, 2009. **30**(16): p. 2785-91.
103. Carapito, R., et al., *Molecular basis of arabinobio-hydrolase activity in phytopathogenic fungi: crystal structure and catalytic mechanism of Fusarium graminearum GH93 exo-alpha-L-arabinanase*. J Biol Chem, 2009. **284**(18): p. 12285-96.
104. de Sanctis, D., et al., *New evidence for the role of calcium in the glycosidase reaction of GH43 arabinanases*. FEBS J, 2010. **277**(21): p. 4562-74.
105. Armougom, F., et al., *Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee*. Nucleic Acids Res, 2006. **34**(Web Server issue): p. W604-8.
106. Taly, J.F., et al., *Using the T-Coffee package to build multiple sequence alignments of protein, RNA, DNA sequences and 3D structures*. Nat Protoc, 2011. **6**(11): p. 1669-82.
107. Quijcho, F.A., Spurlino, J. C., Rodseth, L. E., *Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/ chemosensory receptor*. Structure, 1997. **5**(8): p. 997-1015.
108. Matsumoto, N., et al., *Crystal structures of open and closed forms of cyclo/maltodextrin-binding protein*. FEBS J, 2009. **276**(11): p. 3008-19.
109. Suzuki, R., et al., *Structural and thermodynamic analyses of solute-binding Protein from Bifidobacterium longum specific for core 1 disaccharide and lacto-N-biose I*. J Biol Chem, 2008. **283**(19): p. 13165-73.
110. Cuneo, M.J., et al., *Structural adaptations that modulate monosaccharide, disaccharide, and trisaccharide specificities in periplasmic maltose-binding proteins*. J Mol Biol, 2009. **389**(1): p. 157-66.
111. Vahedi-Faridi, A., et al., *Crystal structures of the solute receptor GacH of Streptomyces glaucescens in complex with acarbose and an acarbose homolog: comparison with the acarbose-loaded maltose-binding protein of Salmonella typhimurium*. J Mol Biol, 2010. **397**(3): p. 709-23.

112. Diez, J., et al., *The crystal structure of a liganded trehalose/maltose-binding protein from the hyperthermophilic Archaeon Thermococcus litoralis at 1.85 Å*. J Mol Biol, 2001. **305**(4): p. 905-15.
113. Evdokimov, A.G., et al., *Structural basis for oligosaccharide recognition by Pyrococcus furiosus maltodextrin-binding protein*. J Mol Biol, 2001. **305**(4): p. 891-904.
114. Cuneo, M.J., et al., *The crystal structure of a thermophilic glucose binding protein reveals adaptations that interconvert mono and di-saccharide binding sites*. J Mol Biol, 2006. **362**(2): p. 259-70.
115. Fielding, L., *NMR Methods for the Determination of Protein-Ligand Dissociation Constants*. Curr Top in Med Chem, 2003. **3**: p. 39-53.
116. Lepre, C.A., Moore, J. M., Peng, J. W., *Theory and Applications of NMR-Based Screening in Pharmaceutical Research*. Chem Rev, 2004. **104**: p. 3641-3674.
117. Angulo, J., P.M. Enriquez-Navas, and P.M. Nieto, *Ligand-receptor binding affinities from saturation transfer difference (STD) NMR spectroscopy: the binding isotherm of STD initial growth rates*. Chemistry, 2010. **16**(26): p. 7803-12.
118. Yan, J., et al., *The effect of relaxation on the epitope mapping by saturation transfer difference NMR*. Journal of Magnetic Resonance, 2003. **163**(2): p. 270-276.
119. Mayer, M., Meyer, B., *Group Epitope Mapping by Saturation Transfer Difference NMR To Identify Segments of a Ligand in Direct Contact with a Protein Receptor*. J Am Chem Soc, 2001. **123**: p. 6108-6117.
120. Mayer, M., James, T. L., *NMR-Based Characterization of Phenothiazines as a RNA Binding Scaffold*. J Am Chem Soc, 2004. **126**: p. 4453-4460.
121. Yuan, Y., et al., *Investigation of binding of UDP-Galf and UDP-[3-F]Galf to UDP-galactopyranose mutase by STD-NMR spectroscopy, molecular dynamics, and CORCEMA-ST calculations*. J Am Chem Soc, 2008. **130**(10): p. 3157-68.
122. Groves, P., Kovér, E. K., André, S., Bendorowicz-Pikula, J., Batta, G., Bruix, M., Buchet, R., Canales, A., Canáda, F. J., Gabius, H-J., Laurents, D. V., Naranjo, J. R., Palczewska, M., Pikula, S., Rial, E., Strzelecka-Kiliszek, A., Jiménez-Barbero, J., *Temperature dependence of ligand-protein complex formation as reflected by saturation transfer difference NMR experiments*. Magn Reson Chem, 2007. **45**: p. 745-748.
123. Spurlino, J.C., Lu, G-Y., Quirocho, F. A. , *The 2.3-Å Resolution Structure of the Maltose-or Maltodextrin-binding Protein, A Primary Receptor of Bacterial Active Transport and Chemotaxis*. J Biol Chem, 1991. **266**(8): p. 5202-19.
124. Viegas, A., et al., *Molecular determinants of ligand specificity in family 11 carbohydrate binding modules: an NMR, X-ray crystallography and computational chemistry approach*. FEBS J, 2008. **275**(10): p. 2524-35.
125. Viegas, A., et al., *Solution structure, dynamics and binding studies of a family 11 carbohydrate-binding module from Clostridium thermocellum (CtCBM11)*. Biochem J, 2013. **451**(2): p. 289-300.

126. Zhang, Y., *I-TASSER server for protein 3D structure prediction*. BMC Bioinformatics, 2008. **9**: p. 40.
127. Zhang, Y. and J. Skolnick, *TM-align: a protein structure alignment algorithm based on the TM-score*. Nucleic Acids Res, 2005. **33**(7): p. 2302-9.
128. Tonozuka, T., et al., *Structural basis for cyclodextrin recognition by Thermoactinomyces vulgaris cyclo/maltodextrin-binding protein*. FEBS J, 2007. **274**(8): p. 2109-20.
129. Duan, X., et al., *Crystal structures of the maltodextrin/maltose-binding protein complexed with reduced oligosaccharides: flexibility of tertiary structure and ligand binding*. J Mol Biol, 2001. **306**(5): p. 1115-26.
130. Kitchen, D.B., et al., *Docking and scoring in virtual screening for drug discovery: methods and applications*. Nat Rev Drug Discov, 2004. **3**(11): p. 935-49.
131. Morris, G.M., Goodsell, D. S., Halliday, R.S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J., *Automated Docking Using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function*. J. Comput. Chem., 1998. **19**: p. 1639-1662.
132. Huey, R., et al., *A semiempirical free energy force field with charge-based desolvation*. J Comput Chem, 2007. **28**(6): p. 1145-52.
133. Quiocho, F.A., Vyas, N. K., *Novel stereospecificity of the L-arabinose-binding protein*. Nature, 1984. **310**: p. 381-386.
134. Gehring, K., Williams, P. G., Pelton, J. G., Morimoto, H., Wemer, D. E. , *Tritium NMR Spectroscopy of ligand binding to Maltose-Binding-Protein*. Biochemistry, 1991: p. 5524-5531.
135. Glynou, K., Ioannou, P. C., Christopoulos, T. K., *One-step purification and refolding of recombinant photoprotein aequorin by immobilized metal-ion affinity chromatography*. Protein Expr Purif, 2003. **27**(2): p. 384-90.
136. Lemercier, G., N. Bakalara, and X. Santarelli, *On-column refolding of an insoluble histidine tag recombinant exopolyphosphatase from Trypanosoma brucei overexpressed in Escherichia coli*. Journal of Chromatography B, 2003. **786**(1-2): p. 305-309.
137. Rehm, B.H.A., Qi, Q., Beermann, Br. B., Hinz, H-J., Steinbuchel, A., *Matrix-assisted in vitro refolding of Pseudomonas aeruginosa class II polyhydroxyalkanoate synthase from inclusion bodies produced in recombinant Escherichia coli*. Biochem J, 2001. **358**: p. 263-268.
138. Schlager, B., A. Straessle, and E. Hafen, *Use of anionic denaturing detergents to purify insoluble proteins after overexpression*. BMC Biotechnol, 2012. **12**: p. 95.
139. Suzuki, H., Terada, T., *Removal of Dodecyl Sulfate from Protein Solution*. Anal Biochem, 1988. **172**: p. 259-263.
140. Lilie, H., Schwarz, E., Rudolph, R., *Advances in refolding of proteins produced in E. coli*. Curr Opin Biotechnol, 1998. **9**: p. 497-501.
141. Tsumoto, K., et al., *Practical considerations in refolding proteins from inclusion bodies*. Protein Expression and Purification, 2003. **28**(1): p. 1-8.
142. Schein, C.H., Noteborn, M. H. M., *Formation of Soluble Recombinant Proteins in Escherichia Coli is Favored by Lower Growth Temperature*. Nature Biotechnology, 1988 **6**: p. 291 - 294

143. Baneyx, F., *Recombinant protein expression in Escherichia Coli*. Curr Opin Biotechnol, 1999. **10**: p. 411-421.
144. Levitt, M.H., *Spin Dynamics, Basics of Nuclear Magnetic Resonance*, 2^o edition. John Wiley & Sons, 2008.
145. Claridge, T.D.W., *High-Resolution NMR Techniques in Organic Chemistry*, Second edition. Elsevier, 2009. **27**.

VI. Apêndice

VI.1. Fundamentos teóricos da técnica de RMN

VI.1.1 Desvio químico

Na ausência de um campo magnético externo B_0 , os momentos angulares de spin de um núcleo em estudo encontram-se distribuídos isotropicamente, isto é a polarização dos spins encontra-se apontada com igual probabilidade para todas as direções. Quando estes spins são colocados na presença de um campo magnético externo B_0 estes precessam em torno do mesmo à frequência de Larmor. Devido essencialmente à presença de pequenos campos magnéticos gerados por correntes elétricas (movimento orbital de electrões), bem como por momentos magnéticos dos núcleos e electrões presentes nas moléculas em estudo, irão desenvolver-se pequenos campos magnéticos locais (B_{loc}).

O campo magnético efetivo (B_{ef}) experimentado por um determinado núcleo está na origem do desvio químico e resulta da soma do campo magnético externo B_0 com o campo magnético local B_{loc} provocado pelos electrões devido a B_0 . A equação 1 representa a relação entre B_{ef} e B_0 em função da constante de blindagem σ que permite contabilizar a contribuição de B_{loc} .

$$B_{ef} = B_0(1 - \sigma) \quad \text{Equação 1}$$

A constante de blindagem σ pode ter contribuições diamagnéticas (se o campo magnético local adoptar a mesma direção de B_0), contribuições paramagnéticas (quando o campo magnético adopta direção oposta à de B_0) ou contribuições locais (equação 2), tais como as zonas de blindagem locais devido ao formação de correntes face ao movimento de electrões deslocalizados em orbitais p em resíduos aromáticos.

$$\sigma = \sigma_{dia} + \sigma_{para} + \sigma_{loc} \quad \text{Equação 2}$$

Assim, alterações no desvio químico de um núcleo de um ligando na presença de uma proteína ou de um núcleo da proteína na presença do ligando, estão em última análise associadas a modificações do ambiente químico local, funcionando como prova de interação molecular entre o ligando e a proteína.

VI.1.2. Relaxação longitudinal T_1 e relaxação transversal T_2

Na ausência de um campo magnético externo B_0 , os momentos angulares de spin de um núcleo em estudo encontram-se distribuídos isotropicamente, isto é a polarização dos spins encontra-se apontada com igual probabilidade para todas as direções. Quando estes spins são colocados na presença de um campo magnético externo (B_0) estes irão começar a precessar em torno do mesmo à frequência de Larmor. Para além de B_0 existem outras fontes de campos magnéticos em estudo tais como: a presença de pequenos campos magnéticos gerados por correntes elétricas (movimento orbital de electrões), bem como por momentos magnéticos dos núcleos e electrões presentes nas moléculas em estudo. Face ao movimento das moléculas estes núcleos irão experienciar campos

magnéticos efetivos resultantes da soma do B_0 e dos pequenos campos magnéticos estes últimos com diferentes magnitudes e orientações em dois intervalos de tempo diferentes.

São estas pequenas flutuações do campo magnético que irão destruir a isotropia da polarização de spins, conduzindo à distribuição anisotrópica de spins no espaço, com a evolução de um momento magnético nuclear global de spins M_0 alinhado com o campo magnético, como se encontra ilustrado na figura VI.1. O equilíbrio térmico é então atingido registando-se um ligeiro excesso populacional de spins num estado de mais baixa energia.

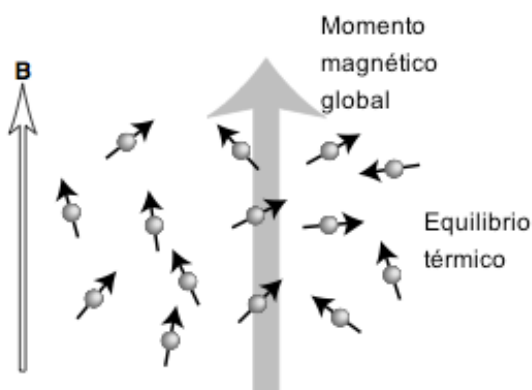


Figura VI.1. Distribuição dos spins nucleares em equilíbrio térmico de Boltzmann. Ligeiro aumento de spins nucleares alinhados com o campo magnético, num estado de mais baixa energia α , após indução de B_0 e formação de dois estados energéticos degenerados. Figura adaptada da referência [144].

Se a polarização global M_0 for perturbada por um pulso de 90° (emissão de radiofrequência perpendicular ao eixo Z), M_0 é transferida para o plano X-Y, e consequentemente a distribuição dos spins nos estados energéticos é alterada, deixando de existir um excesso de spins paralelos com B_0 . Neste caso irá ocorrer um igualar de populações distribuídas entre os níveis energéticos α e β . A relaxação T_1 é o processo através do qual os spins irão perder energia para o ambiente envolvente (relaxação *spin-lattice*) voltando ao equilíbrio térmico. Para que ocorra relaxação T_1 é necessária a presença de um campo magnético oscilante com uma energia igual à diferença de energia entre os dois estados energéticos, por outras palavras é necessário um campo magnético cuja frequência seja igual à frequência de precessão dos núcleos- frequência de Larmor. A relaxação T_1 pode ser mais ou menos estimulada, consoante o regime de movimento da molécula em estudo, como é descrito nos capítulos posteriores (V.1.4. *Movimento molecular e relaxação de spins*, bem como no capítulo V.1.5. *Tempo de correlação e Densidade espectral*).

Quando a polarização dos spins é “enviada” para o plano X-Y, esta começa a precessar ao longo deste plano à frequência de Larmor. Como foi dito em cima, à escala microscópica cada spin irá “sentir” nas suas vizinhanças um campo magnético efetivo, não sendo este exatamente o mesmo para todos os spins presentes na molécula. Assim, existirão núcleos que irão sentir um campo magnético mais intenso e consequentemente irão precessar com uma frequência mais elevada, e haverá um conjunto de spins que irão sentir um campo magnético menos intenso e irão precessar a frequências mais baixas. Num instante inicial M_0 está no plano XY sendo que os spins encontram-se em coerência, no entanto face às diferentes frequências de precessão dos spins, a sua distribuição ao longo do plano irá ficar isotrópica após um intervalo de tempo, com a consequente eliminação do vector M_0 , e perda de coerência do sistema. Este tipo de relaxação é denominada de relaxação

transversal T_2 . Para além dos campos magnéticos locais, se o campo magnético externo não for homogêneo em toda a amostra, tal como foi explicado em cima os spins que “sentirem” mais campo irão precessar mais rapidamente e vice-versa, levando à destruição da magnetização, podendo o campo magnético ser homogeneizado na amostra através do *shimming* da amostra.

Associado ao T_2 (equação 3) está a largura da Lorentziana de absorção, como é demonstrado pela equação, onde T_2^* é o somatório da i) heterogeneidade de B_0 e ii) campos magnéticos locais.

$$\nu_{1/2} = \frac{1}{\pi T_2^*} \quad \text{Equação 3}$$

A monitorização da largura de linha, pode ser utilizada para estudar a interação ligando-receptor. Após a reconhecimento do ligando este muda o seu regime de rotação, sofrendo consequentemente um alargamento de linha.

Pequenas moléculas apresentam T_2 longo e consequentemente apresentam um largura de linha pequena, ao passo que moléculas grandes (proteínas) irão ter um T_2 curto e uma largura de linha maior. O alargamento de linha está associado ao facto dos diferentes vectores apresentarem uma largo espectro de frequências de precessão.

A relaxação por mecanismo em T_1 responsável pela recuperação da magnetização em Z é também responsável pela perda de magnetização no plano xy. Logo mecanismos que levem a T_1 irão em última análise provocar relaxação T_2 , sendo $T_1 \geq T_2$.

VI.1.3. Efeito Nuclear de Overhauser (NOE)

O NOE pode ser definido como a mudança de intensidade de uma ressonância, quando as populações de um núcleo adjacente são perturbadas do seu equilíbrio térmico. A magnitude pode ser expressa enquanto uma diferença relativa de intensidades no equilíbrio I_0 , e na presença do efeito de NOE, I.

$$NOE_I^S = \frac{I - I_0}{I_0} \quad \text{Equação 4}$$

A equação 4 representa o efeito NOE sentido pelo núcleo I quando o núcleo S é perturbado, sendo I e I_0 a intensidade do sinal na presença e ausência de irradiação do núcleo S. O efeito de NOE sentido pode ser positivo ou negativo, sendo este comportamento ditado pelo movimento molecular dos núcleos em solução (ver capítulo, VI.1.5 *Tempo de correlação e função densidade espectral*).

Para melhor elucidar o efeito NOE, podemos considerar um sistema homonuclear de dois núcleos I e S de spin $1/2$. Estes dois núcleos não partilham acoplamento escalar ($J_{IS}=0$), isto é, não apresentam ligações químicas entre si, no entanto encontram-se suficientemente próximos espacialmente para que ocorra entre ambos acoplamento dipolar (acoplamento espacial entre dois spins nucleares através da sua componente magnética).

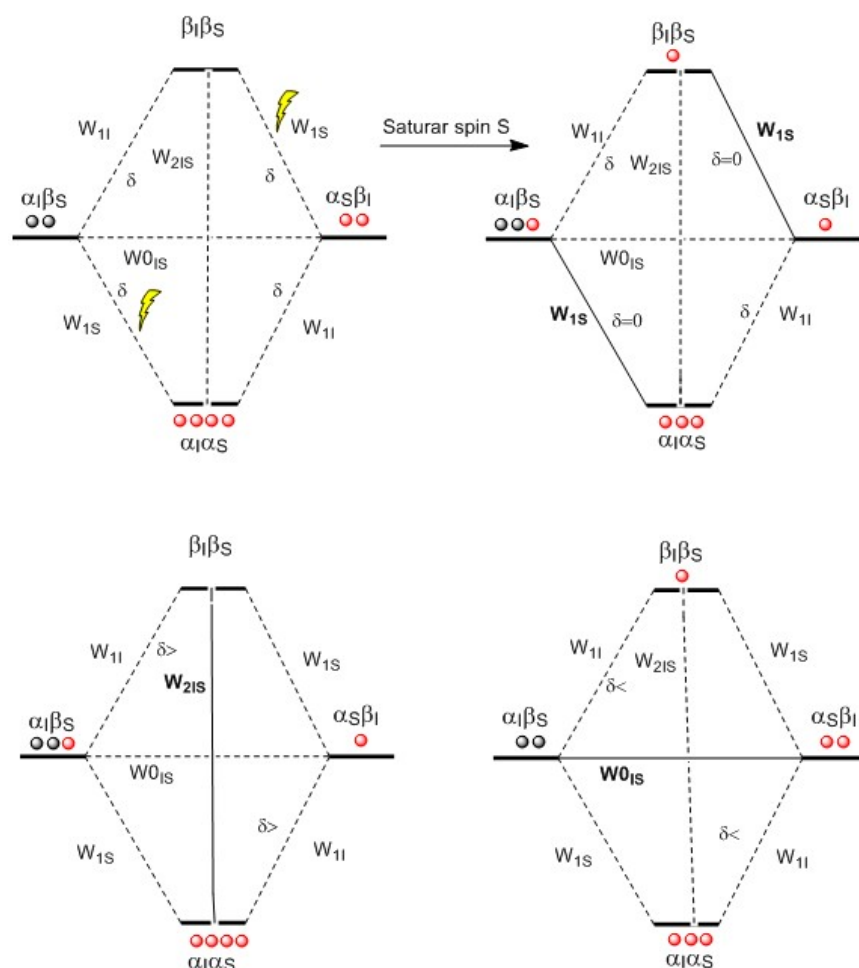


Figura VI.2. Efeito NOE. Na parte superior é possível observar a saturação dos spins nucleares S. Na parte inferior, é possível observar à esquerda a formação de NOE positivos, ao passo que à esquerda NOE negativo, ambos para o núcleo I.

Na parte superior da figura VI.2 é possível observar que ocorre uma saturação dos spins nucleares S, ou seja a diferença populacional nos dois estados energéticos (α_S e β_S) passou a ser nula. Através da relaxação longitudinal dos spins, o sistema irá procurar atingir um novo equilíbrio térmico. As vias através das quais irá ocorrer a relaxação de spin são:

1. *Transições double quantum* (W_{2IS}) que representam a probabilidade de ocorrer transições entre níveis energéticos $\beta\beta$ - $\alpha\alpha$. Com o intuito de restabelecer o equilíbrio térmico das populações dos spins nucleares S, os spins nucleares nos estados $\beta\beta$ irão relaxar para $\alpha\alpha$, promovendo um aumento da diferença entre as populações I (α_I e β_I), e consequentemente registrar-se-à um aumento da intensidade da ressonância do núcleo I. Este fenómeno é denominado de NOE positivos (ver capítulo, VI.4 *Movimento molecular e relaxação de spin*).
2. *Transições de zero quantum* (W_{0IS}), que descrevem a probabilidade de ocorrer transições entre níveis energéticos $\alpha_I\beta_S$ e $\alpha_S\beta_I$. Analogamente ao mecanismo descrito anteriormente, haverá relaxação de spin de $\alpha_S\beta_I$ para $\alpha_I\beta_S$, promovendo consequente uma diminuição das diferença de populações entre os diferentes estados energéticos afetos aos spins nucleares I. Este fenómeno irá promover uma diminuição da intensidade da ressonância associada a esse núcleo- NOE negativo (ver capítulo, VI.4 *Movimento molecular e relaxação de spin*).

$$NOE_I^S = \frac{\gamma_S}{\gamma_I} \left[\frac{W_2 - W_0}{W_0 + 2W_1^I + W_2} \right] = \frac{\gamma_S}{\gamma_I} \left[\frac{\sigma_{IS}}{\rho_{IS}} \right]$$

Equação 5

A equação 5 resume o que até agora foi estabelecido. O sinal do NOE irá depender da eficiência dos processos de relaxação denominados *cross-relaxation* (σ_{IS}). Quando os processos de relaxação através das transições W_2 prevalecem a função NOE_I^S assume magnitude positiva e quando a relaxação ocorre por transições W_0 a função NOE_I^S assume magnitude negativa. É de referir que a transição W_1^S encontra-se saturada, não podendo ocorrer relaxação por esta via, no entanto a via de relaxação W_1^I encontra-se livre podendo “diluir” ou até mesmo abolir o efeito NOE.

VI.1.4. Movimento molecular e relaxação de spin

Como foi explicado anteriormente, o NOE é sentido após os spins nucleares relaxarem por vias que competem entre si (*cross-relaxation pathways*), originando NOE's positivos e negativos, consoante a reorganização dos spins nucleares. Após a aplicação de uma radiofrequência (à frequência de Larmor) oscilante perpendicularmente à magnetização M_0 - pulso de 90° - o sistema é perturbado do seu equilíbrio térmico, no qual há um ligeiro excesso de populações no estado α de mais baixa energia, passando a magnetização para o plano transversal, sendo que neste ponto há um numero igual de populações no estado α e β , como é descrito na figura VI.3.

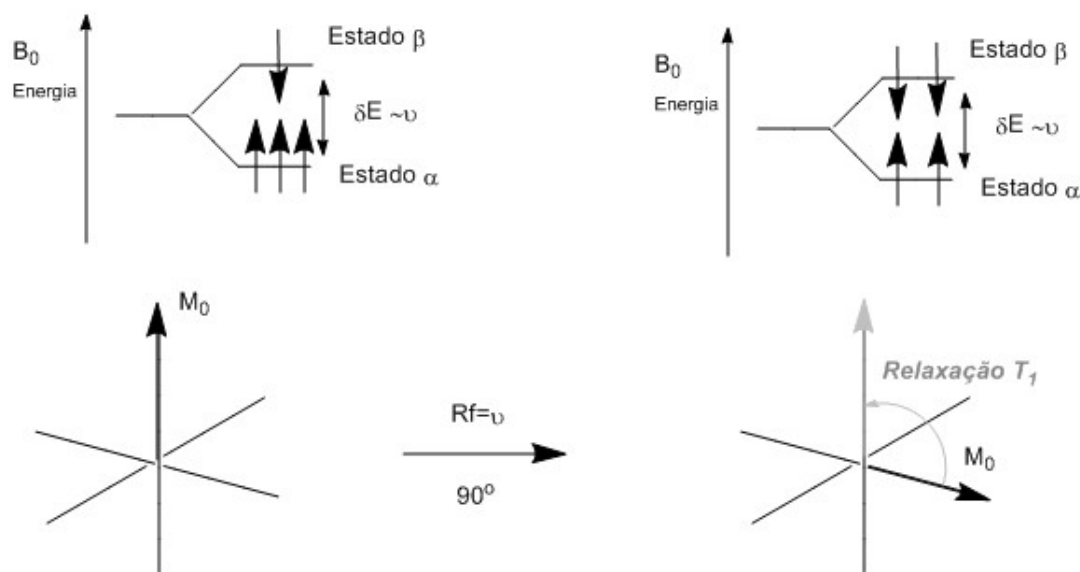


Figura VI.3. Manipulação de spins nucleares após aplicação de rf em ressonância. Transições populacionais dos spins nucleares estão na base da origem do fenómeno de RMN. A redistribuição populacional apenas é conseguida na presença de um campo magnético flutuante (B_1) cuja frequência é igual à frequência Larmor (ν). M_0 representa a média dos momentos magnéticos associados aos spins nucleares do sistema.

Portanto, a aplicação de uma radiofrequência com um componente magnética oscilatória à frequência de Larmor irá promover transições populacionais. No entanto, estas transições energéticas podem também elas acontecer fruto do movimento molecular. Neste caso, para efeitos de NOE em núcleos de spin $1/2$ os principais processos que estão na origem da relaxação dos spins são exercidos pelos pequenos campos magnéticos inerentes de cada núcleo. Dois núcleos acoplados dipolarmente irão experimentar o campo magnético do núcleo vizinho para além do campo magnético

externo. O movimento molecular irá desempenhar um papel por um lado na magnitude e direção destes campos magnéticos (microscópicos), sendo que a frequência deste movimento irá gerar um campo magnético oscilatório à frequência de Larmor responsável pela relaxação longitudinal dipolar.

VI.1.5. Tempo de correlação e densidade espectral

No entanto nem todos os movimentos moleculares irão originar campos magnéticos oscilatórios à frequência exata de Larmor necessária para induzir a transição populacional dos spins entre os níveis energéticos α e β .

O movimento molecular pode ser expresso pelo tempo de correlação (Γ_c) que traduz o tempo que a molécula demora a efetuar uma rotação de 1 radiano. Assim moléculas com um movimento de rotação rápido apresentam um tempo de correlação curto, ao passo que moléculas cuja a rotação é lenta apresentam um tempo de correlação elevado. Sendo o tempo de correlação proporcional ao peso molecular da molécula. Em conclusão moléculas de elevado peso molecular têm um tempo de correlação elevado (nanossegundos) e consequentemente um movimento molecular em solução lento ao passo que moléculas pequenas irão apresentar um tempo de rotação na ordem dos picossegundos, tendo um movimento molecular em solução rápido.

A função densidade espectral, ilustrada na figura VI.4, pode ser vista como a probabilidade que um movimento molecular tem em gerar um campo magnético oscilante à frequência de Larmor.

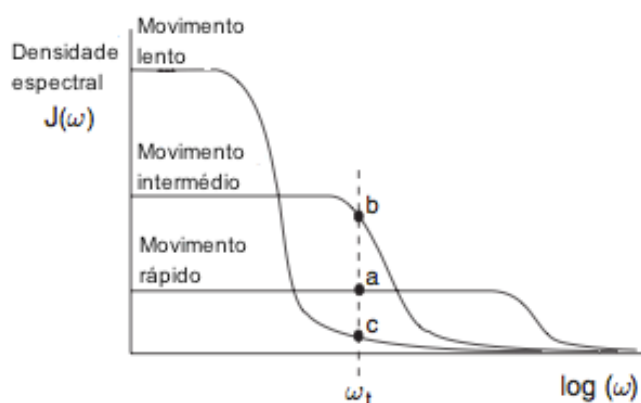


Figura VI.4. Função densidade espectral para diferentes regimes de movimento molecular. Figura adaptada da referencia [145].

Como é possível observar na figura acima, moléculas pequenas apresentam uma igual probabilidade de gerar campos magnéticos oscilantes de baixas e altas frequências. Ao passo que moléculas de elevado peso molecular como é o caso das proteínas, apresentam uma probabilidade baixa de gerar campos magnéticos oscilantes à frequência de Larmor, sendo que apresentam uma elevada probabilidade de gerar campos magnéticos de baixas frequências. Esta análise permite prever como é que as vias de relaxação irão depender do tempo de correlação.

As transições W_0 encontram-se associadas a frequências de Larmor baixas na ordem os kHz/Hz ($\omega_L = \omega_I - \omega_S$). Logo moléculas com regimes de movimento molecular lento irão promover a relaxação de spins por esta via. Por outro lado, moléculas de pequenas dimensões são as que maior probabilidade apresentam de induzir transições energéticas de elevadas frequências como são o caso das transições *Double Quantum* $W_2 = \omega_I + \omega_S = 2\omega_L$. Assim para moléculas de maiores dimensões

(proteínas), as vias de relaxação privilegiadas são as transições de *Zero Quantum*, originando NOE de sinal negativo. Por outro lado moléculas de menores dimensões irão privilegiar as transições de *Double Quantum* W_2 e apresentam NOE positivos.

VI.2. Meios de cultura

Tabela VI.1. Preparação de meio Luria-Bertani (LB).

Componente	Quantidade por 1L (g)
Extrato de levedura	5
NaCl	10
Triptona	10

Tabela VI.2. Preparação de meio mínimo M9.

Componente	Quantidade
Solução M9	100mL/L
Glucose	0,4% (p/v)
MgSO ₄ .7H ₂ O	2mM
Tiamina	1mL/L de solução 1mg/L

Tabela VI.3. Preparação de solução de sais M9.

Componente	Quantidade
¹⁵ NH ₄ Cl	1g/L
KH ₂ PO ₄	60g/L
Na ₂ KPO ₄	120g/L

Tabela VI.4. Preparação de meio auto indução

Componentes de solução stock 50X5052	Quantidade
Glicerol	25% (p/v)
D-Glucose	2,5% (p/v)
α- Lactose	10% (p/v)

Componentes de solução stock 20XNPS	Quantidade
(NH ₄) ₂ SO ₄	500mM
KH ₂ PO ₄	1M
Na ₂ PO ₄	1M

Componente	Quantidade
MgSO ₄	1:1000
50X5052	1:50
20XNPS	1:20
Meio LB	92,9 mL

VI.3. Quantificação pelo método de Bradford

Tabela VI.5. Tabela para a preparação de amostras para a construção da recta de calibração

[BSA] mg/mL	Abs _{595nm}	σ
0,01	0,063	0,005
0,02	0,119	0,002
0,04	0,249	0,004
0,06	0,336	0,009
0,08	0,428	0,011
0,10	0,537	0,019

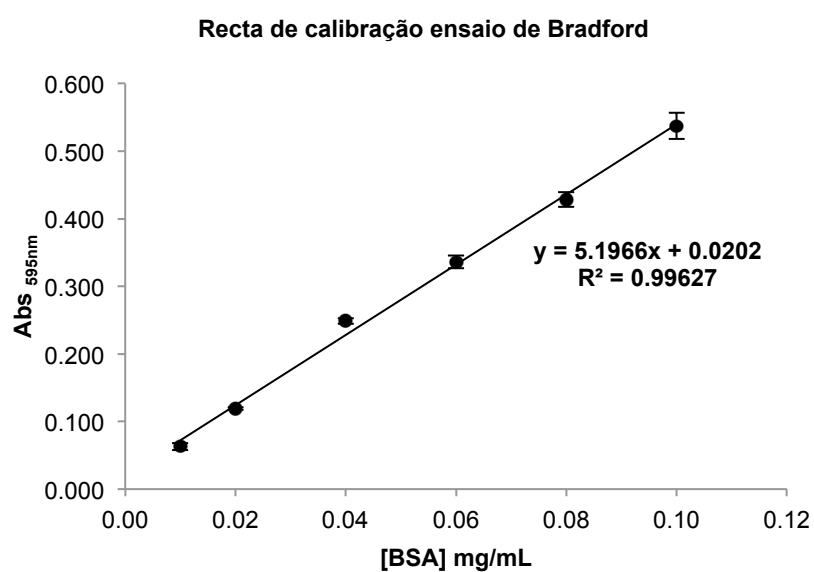


Figura VI.5. Recta de calibração do método de Bradford. ($y=5.1966x + 0.0202$).

VI.4 Dados de NMR

Tabela VI.6. Dados experimentais. Valores de A_{STD} experimentais e calculados (a), para diferentes tempos de saturação. São apresentados também os parâmetros de STD_{Max} , K_{sat} , STD_{fit} , STD_{fit} normalizado.

Epítipo	0,10	0,20	0,40	0,60	0,80	1,00	1,20	1,40	1,60	1,80	2,00	2,50	3,00	3,50	STD Max	Ksat	STD fit	STDfit %
H4a+H4m	0,76	1,52	1,15	0,78	1,99	2,47	2,19	2,62	2,72	3,13	3,14	3,51	3,84	3,95	4,26	0,71	3,00	100,0
H4a+ H4m(a)	0,29	0,56	1,05	1,47	1,84	2,15	2,43	2,67	2,88	3,06	3,22	3,53	3,74	3,90				
H2m+H2n	0,40	0,71	0,76	0,64	1,44	1,78	1,94	2,24	2,51	2,91	3,00	3,47	3,67	4,08	5,75	0,36	2,05	68,36
H2m+H2n(a)	0,20	0,40	0,77	1,11	1,43	1,73	2,00	2,26	2,50	2,73	2,93	3,40	3,78	4,10				
H4n+H2β+H3β	0,32	0,68	0,65	0,59	1,26	1,58	1,64	1,85	2,04	2,24	2,37	2,62	2,81	2,97	3,54	0,53	1,89	63,00
H4n+H2β+H3β(a)	0,18	0,36	0,68	0,97	1,23	1,47	1,68	1,86	2,03	2,19	2,32	2,61	2,83	2,99				
H2α+H3α+H3m	0,38	0,69	0,70	0,64	1,36	1,54	1,62	1,76	1,91	2,12	2,26	2,48	2,51	2,63	2,91	0,70	2,04	67,97
H2α+H3α+H3m(a)	0,20	0,38	0,71	1,00	1,25	1,47	1,66	1,82	1,96	2,09	2,19	2,40	2,55	2,66				
H3n+H4β	0,41	0,79	0,82	0,72	1,48	1,69	1,84	2,08	2,23	2,49	2,58	2,89	3,11	3,25	3,77	0,58	2,17	72,26
H3n+H4β(a)	0,21	0,41	0,78	1,10	1,39	1,65	1,88	2,09	2,27	2,43	2,58	2,88	3,10	3,27				
H5m	0,30	0,49	0,64	0,47	1,06	1,10	1,21	1,34	1,46	1,68	1,79	1,93	1,97	1,99	2,26	0,69	1,56	51,95
H5m(a)	0,15	0,29	0,55	0,77	0,96	1,13	1,27	1,40	1,51	1,61	1,69	1,86	1,97	2,06				
H5α+H5β	0,22	0,42	0,55	0,47	0,96	1,09	1,19	1,26	1,38	1,57	1,59	1,79	1,78	1,92	2,10	0,70	1,46	48,67
H5α+H5β(a)	0,14	0,27	0,51	0,72	0,90	1,05	1,19	1,31	1,41	1,50	1,58	1,73	1,84	1,92				
H5m'	0,19	0,33	0,39	0,31	0,81	0,97	0,94	1,03	1,19	1,31	1,38	1,52	1,51	1,62	1,85	0,63	1,17	38,91
H5m'(a)	0,11	0,22	0,41	0,58	0,73	0,87	0,98	1,09	1,18	1,26	1,33	1,47	1,57	1,65				
H5n	0,13	0,28	0,35	0,31	0,72	0,88	0,90	1,01	1,16	1,25	1,36	1,46	1,48	1,58	1,86	0,58	1,08	35,81
H5n(a)	0,10	0,20	0,38	0,55	0,69	0,82	0,93	1,03	1,12	1,20	1,28	1,42	1,53	1,61				
H5α'+H5β'	0,21	0,42	0,53	0,29	0,97	1,07	0,94	1,07	1,28	1,54	1,36	1,46	1,65	1,36	1,63	0,90	1,47	48,85
H5α'+H5β'(a)	0,14	0,27	0,49	0,68	0,84	0,97	1,08	1,17	1,24	1,31	1,36	1,46	1,52	1,56				
H5n'	0,29	0,59	0,81	0,66	1,30	1,54	1,54	1,79	2,00	2,29	2,36	2,60	2,80	2,79	3,32	0,59	1,95	64,89
H5n(a)'	0,19	0,37	0,69	0,99	1,24	1,47	1,68	1,86	2,02	2,17	2,29	2,55	2,75	2,89				

VI.5 Espectros suplementares

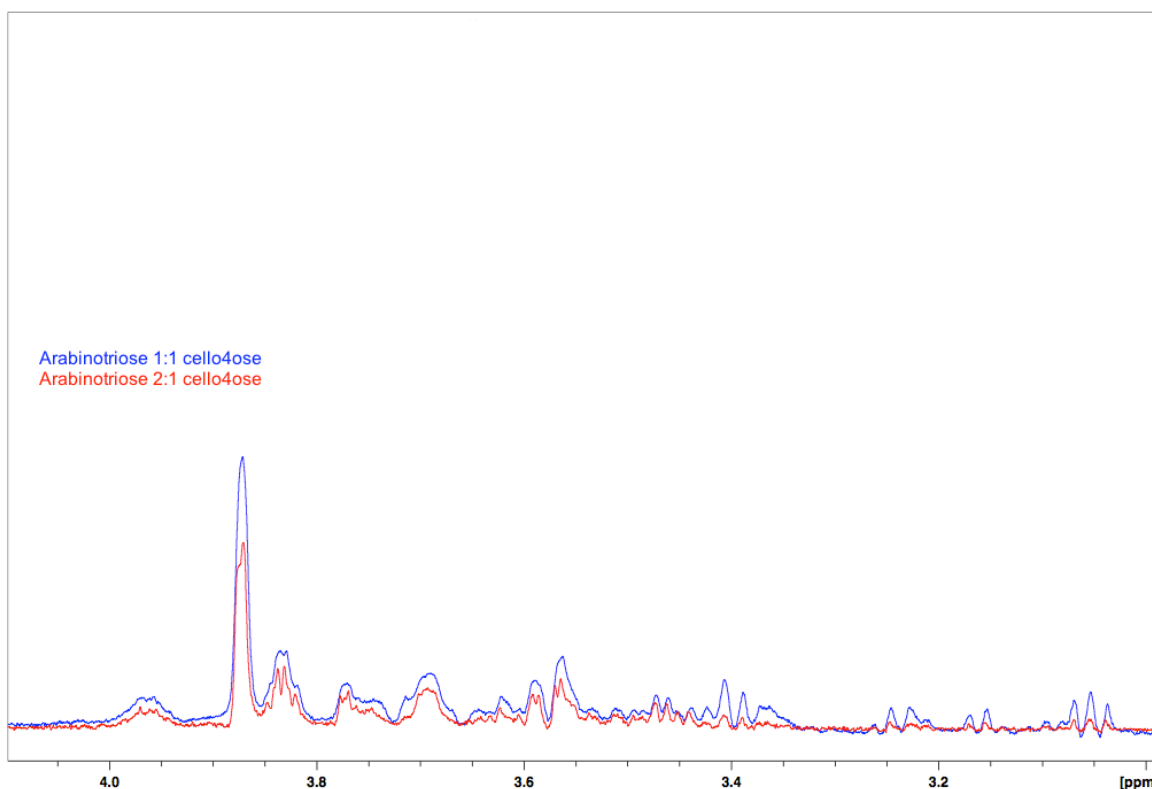


Figura VI.6. Espectro de STD-RMN em condições de competição da arabinotriose/celotetraose com AraN. Espectro ^1H -STD (I_{STD}) a azul representa o competição arabinotriose 1:1 celotetraose. Espectro ^1H -STD (I_{STD}) a vermelho representa o espectro arabinotriose 2:1 celotetraose.

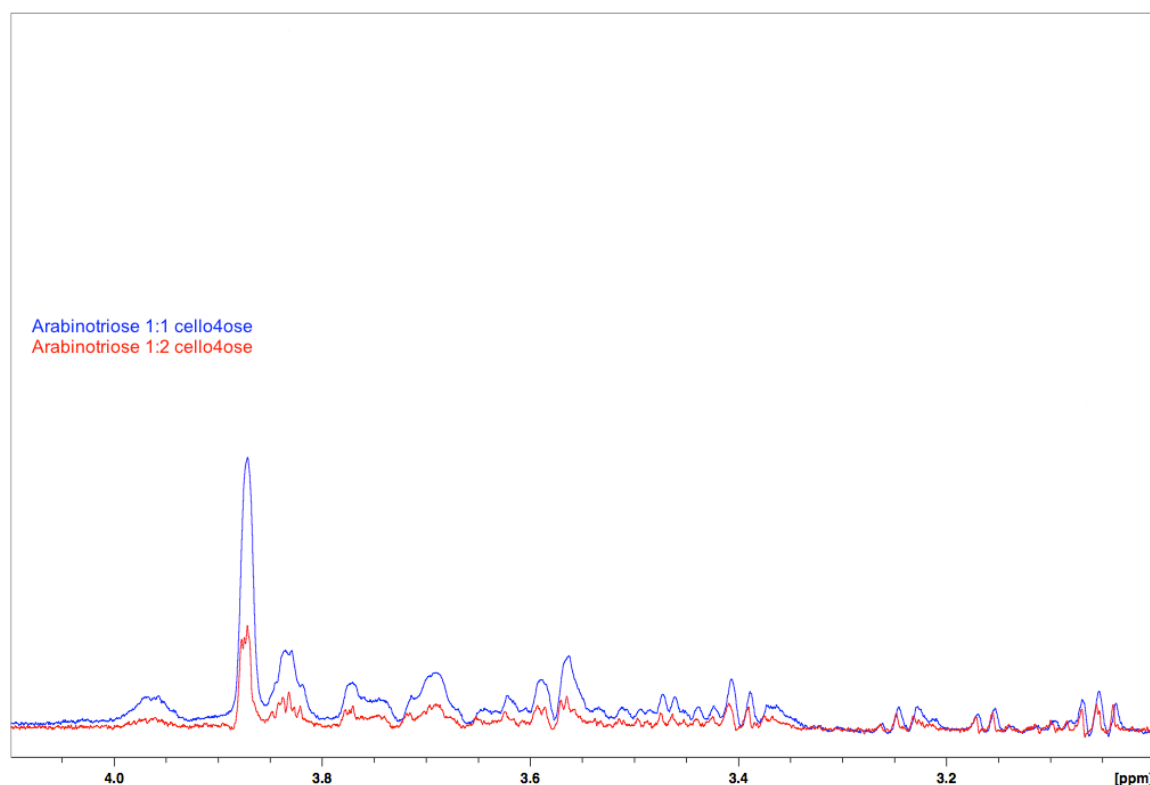


Figura VI.7. Espectro de STD-RMN em condições de competição da arabinotriose/celotetraose com AraN. Espectro ^1H -STD (I_{STD}) a azul representa o competição arabinotriose 1:1 celotetraose. Espectro ^1H -STD (I_{STD}) a vermelho representa o espectro arabinotriose 1:2 celotetraose.